DOCUMENT RESUME

ED 041 490 FL 001 163

AUTHOR	Agard, Frederick B.; Dunkel, Harold B.
TITLE	An Investigation of Second-Language Teaching.
PUB DATE	48
NOTE	349p.
EDRS PRICE	EDRS Price MF-$1.50 HC Not Available from EDRS.
DESCRIPTORS	Advanced Programs, *Audiolingual Skills, Colleges, Course Content, Educational Objectives, English (Second Language), *Language Instruction, Language Proficiency, *Language Programs, *Language Research, Language Tests, Program Descriptions, Secondary Schools, Second Language Learning, Statistical Data, *Testing Programs

ABSTRACT

This investigation of high school and college language programs in the 1940's presents information on the objectives of language teaching and the measurement of language proficiency. An aural-oral testing program is described, and the aural-oral experimental groups are compared with conventionally taught classes. Advanced levels of language skill are studied in programs teaching English as a foreign language. An appendix includes additional data. For a companion document see ED 037 104. [Not available in hard copy due to marginal legibility of original document.] (AF)

An Investigation of Second-Language Teaching

By Frederick B. Agard
Cornell University

and Harold B. Dunkel
The University of Chicago

ED0 41490

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FL 001 163

Ginn and Company

Boston - New York - Chicago - Atlanta - Dallas - Columbus
San Francisco - Toronto - London

Copyright, 1948, by Ginn and Company
All Rights Reserved

748.12

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Ginn and Company TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Preface

The studies contained in the two volumes of this report (AN INVESTIGATION OF SECOND-LANGUAGE TEACHING and SECOND-LANGUAGE LEARNING) were possible only through the help we received from many quarters.

Scores of language teachers supplied the data contained in the following pages. In some instances our demands on them were slight, requiring only the administration of the tests and the return of the data. In other cases our collaborators, especially the directors of experimental programs, devoted many hours to our interests, preparing statements about their programs, gathering supplementary data for us, and otherwise making possible a more complete account. It would be impossible to attempt to distribute appropriate specific acknowledgements over so wide a range of indebtedness. We can only express our obligation and thanks to all collectively.

During the investigation, Edna Davis, Yolanda Leite, and Hazel Mitchell worked for long periods with the staff in Chicago, and Alice Ann Chambers and Einar Ryden worked at Northwestern University. In addition to the acknowledgements of their work made in the appropriate places throughout the volumes, we wish to thank them here for their efforts. The same is true of research assistants who worked for shorter periods at the various experimental centers

We owe a particularly great debt to Lucas T. Clarkston for the conscientious care with which he scored the masses of tests and carried out the statistical analyses.

FL 001 163

Our colleagues Ralph W. Tyler, the director, and Hugh R. Walpole, have, of course, contributed at every stage of the project; and their suggestions, both major and minor, in regard to these reports have done much to improve them. Needless to say, our colleagues cannot be held responsible for the statements and conclusions we advance here.

Finally, the whole undertaking was possible only because of a generous grant by the Rockefeller Foundation to the University of Chicago.

Though we divided the responsibility for preparing the original draft of this volume, we have subsequently worked over all the material together, and the present statement represents the views of us both.

We are grateful to Henry Holt and Co. for permission to quote from Leonard Bloomfield's *Language* and to The University of Michigan Press for similar permission regarding C. C. Fries' *Teaching and Learning English as a Foreign Language.*

Frederick B. Agard
Harold B. Dunkel

Contents

Preface iii

CHAPTER

I · The Purpose of the Investigation 1
- Desiderata for an Adequate Study 5
- The Present Investigation 8

II · The Objectives of Language Teaching 15
- Reasons for Earlier Minimization of the Oral-Aural Skills 18
- Recent Arguments for Greater Emphasis on the Oral-Aural Skills 22
- "Knowledge" of a Language 25
- The Problem in Terms of Vocabulary Selection 28
- Defining the Oral-Aural Skills 36

III · The Measurement of Language Proficiency 39
- Reading Tests 40
- Tests of Aural Comprehension 41
  - Problems of construction 45
  - Problems of reliability 49
  - Evaluation of the aural tests 51
- Tests of Oral Production 55
- Tests of Writing 60

IV · The Aural and Oral Testing Program 61
- Aural Comprehension Tests 62
  - Standardization of the aural tests 62
  - College norms 64

Secondary-School norms 74
Evaluation of general aural achievement 80
Oral Production Tests 97
Samples of oral production 100
Evaluation of oral performance 115
Summary 116
V · Conventional Groups Not Stressing Aural-Oral Objectives 117
College A 118
College B 134
College C 141
College D 146
College E 148
College F 150
College G 151
High School A 154
Summary 157
VI · Experimental Groups Stressing Aural-Oral Skills as an Initial Objective and Introduction to Reading 159
College H 160
College J 177
College B 207
College C 216
College K 219
College L 224
College M 230
College N 236
High School B 240
High School C 243
High School D 245
VII · The Advanced Level of Language Skill: English as a Foreign Language 247
Complicating Factors 249
Construction of the Tests 254

The Elementary Structure Test 255
Reading Comprehension 256
Aural Comprehension 258
Writing 263
Problems at the Advanced Level 271
Reliabilities and Intercorrelations of the Tests 278

VIII · A Statement of Conclusions and Further Needs 279
General Description of Experimental Courses 280
Achievement in the Experimental Programs 286
Vindication of Assumptions 291
Adherence to Principles 293
Further Needs 294

APPENDIX

A · Norms for the Aural Tests and for The Russian Reading-Vocabulary-Grammar Test 303
B · Charts of Co-operative and Aural Test Data 317
C · Reliabilities and Internal Correlations of the Aural Tests 339

Chapter I

The Purpose of the Investigation

The time at which the Investigation of the Teaching of a Second Language was established, early in 1944, was a very significant one in the recent history of language teaching. The Modern Foreign Language Study of the 1920's had given great impetus to further activity, both on the part of those who accepted the recommendations of the "Coleman Report" and on the part of those who rejected it.

The Second World War further vitalized language teaching. When hostilities were still only imminent, it was obvious that America was linguistically unprepared for a global war. To remedy this defect, the American Council of Learned Societies, with the co-operation of the Linguistic Society of America, set up a program of intensive language courses. Though these were devoted in large part to those languages which had heretofore been little studied, workers in this program were convinced that the general procedure had important implications for the study of all languages, including those which had been taught in our schools for many years. Intensive courses were, consequently, offered in the familiar languages of western Europe as well as in the more exotic tongues.

After America entered the war, the Army Specialized Training Program (ASTP) in language-and-area and other schools established by various governmental departments and agencies carried on similar work. These developments have already been reported in detail elsewhere.[1] Here we need only to sketch the important influence which all this activity, particularly the ASTP, had on language teaching generally. Because the ASTP was carried on in more than fifty colleges and universities, hundreds of language instructors taught in it, and many younger language teachers studied under the system while serving in the armed forces. One of the most stimulating factors in ASTP was that though the Army suggested a general methodology, much was left to the ingenuity and resourcefulness of local staffs. Within the broad limits of the Army directives, there was great latitude for experiment and innovation, an opportunity of which many language staffs took advantage. Even those requirements imposed by the directives stimulated language teaching because they created a new situation. To mention two aspects of this change, most language teaching had aimed primarily at a reading knowledge of the language and had occupied three to five class hours per week of the student's time. The Army, on the contrary, was interested almost exclusively in having the student able to speak the language and to understand it when spoken to him; and for this training it provided approximately seventeen hours weekly. It is no wonder, then, that this almost revolutionary setting, along with the suggestions contained in the recommended methodology, set many language teachers to rethinking familiar problems. Many of them concluded that certain elements of the Army program should be transferred to civilian instruction; and at the time the Investigation began, many colleges were planning more or less elaborate efforts to apply the experience gained through ASTP to peacetime teaching as soon as it could be resumed.

---

[1] R. J. Matthew, *Language and Area Studies in the Armed Services*. Washington: The American Council on Education, 1947. P. A. Angiolillo, *Armed Forces' Foreign Language Teaching*. New York: S. F. Vanni, 1947.

The military programs were not the only force vitalizing language teaching, though their size probably made them the best known and most influential. In the Southwest, the increase in the Spanish-speaking population had confronted several states with a problem of communication among their citizens; and they had accepted the challenge of teaching both Spanish and English as second languages in the elementary schools. These attempts at a lower educational level than was commonly the scene of language teaching had raised new questions and had resulted in experiments with materials, methods, and devices. The teaching of English as a foreign language had also begun to come into its own after being treated as something of a stepchild for many years.[2]

These are only a few of the factors which were producing changes in language teaching at the time our Investigation began. It seemed a likely moment for an attempt to summarize the experience gained from all these enterprises — not a historical account, but rather an analysis of the general principles and specific procedures which would represent what had been learned. As was intellectually healthy, considerable controversy had arisen about these programs. Some believed that the experience derived from these courses ought to revolutionize all language instruction. Others felt that these special programs had been carried on in unusual circumstances which would not be reproduced elsewhere and hence that any new insights, however sound, would not be generally useful. Still others were inclined to echo the famous book review and to insist that the new things in these programs were not good and the good things not new. Unfortunately this controversy could not be resolved by evidence. For various reasons few objective data were available regarding the accomplishments of ASTP, and the personal estimates of different observers varied. Likewise the teaching of English had been carried on largely within the confines of a particular institution or system of instruction, and no comparative evidence from impartial sources existed.

---

[2] Further details will be found in Chapter VII.

This situation could, however, be remedied. At the time the Investigation began, there was an unprecedented interest in language teaching and language learning, an interest which even spilled over into the popular periodicals. With the end of the war in sight, many teachers and colleges were planning experiments in modifying and adapting the techniques of ASTP to civilian instruction. The tide of foreigners seeking to learn English here or in their own countries was rising ever higher. A time when so many unusual language programs would be in action appeared a good one in which to attempt to secure objective evidence on some of the questions which had been raised.

This timing of the Investigation affected the nature of the study and its findings. Starting work in 1944 as the ASTP was about to end, we were able to follow the new developments and experiments as soon as they were instituted in the academic year 1944-1945. More work was possible in 1945-1946 as more colleges and other centers of language teaching recovered from their wartime alterations and put new programs into operation. This was the heyday of language experimentation. In 1946-1947, as the horde of postwar students descended upon the colleges, many of the experimental programs (which made heavy demands on staff, classroom space, and other overtaxed facilities) had to be abandoned.

Yet the period during which we worked also had obvious disadvantages. The wartime and postwar difficulties in securing staff, materials, clerical help, transportation, and nearly everything else, immeasurably hindered the work of both the Investigation and the teaching programs. So, for example, railroad and coal strikes twice coincided with important testing periods, and the consequent restrictions and delays deprived us of irreplaceable data. More important was the fact that we covered experiments in their very early stages. Even if the materials and procedures developed for the ASTP did not need much modification before they could be used with ordinary high school and college students (a point about which there was considerable debate), time is required for organizing any new program and getting it "shaken down." We may have examined experimental foreign-language courses too soon to get a fair picture of what is ultimately possible, but in most cases this early period is the only one for which evidence is available. Possibly the first impetus and enthusiasm which attended these undertakings made this handicap less severe than might be expected. We can, at any rate, report what was accomplished.

DESIDERATA FOR AN ADEQUATE STUDY

The circumstances under which the Investigation operated suggested rather clearly the problems about which its work should center and hence the general outline of this report. Possibly the best way to set forth what was involved is to begin by stating the desiderata for a technically adequate study. We can then indicate the points at which our present investigation meets or falls short of these requirements. The shortcomings are particularly important. For one thing they indicate the reservations which must be kept in mind in examining the data we secured. Of still greater importance is the fact that many of these difficulties can be overcome only by the co-operative effort of all language teachers and hence a more definitive study is impossible until more work has been done in certain areas.

I. Adequate Coverage

A language investigation should cover a wide number of cases. Only by such broad coverage can it be certain of surveying the full range of possibilities offered by varied or conflicting hypotheses. Extended exploration can also make certain that the best and the most advanced programs have been covered as well as the run-of-the-mill variety. Both sorts of information are important. While one should know what can be achieved by the best programs under the most favorable circumstances, one should not lose sight of the accomplishment of language teaching as a whole and assume that these outstanding examples are typical of what is or can be done everywhere.

II. A Clear Statement of Objectives

Only in comparison with the aims which language teaching seeks can the degree of its achievement or the effectiveness of its procedures be judged. The objectives must be stated with such clarity and specificity that the attainment of them can be accurately measured and that general agreement can be reached, not only concerning the nature of the objectives sought, but also in regard to success or failure in achieving them. As for specificity, statements of objectives must find some middle ground between two extremes. "Teaching the student to speak Spanish" is, on the one hand, too general to be useful. We must have a more precise definition of what degree of oral command is meant. On the other hand, stating that certain materials must be covered or a certain test must be passed leaves the objectives too narrow or particularized. There must be some clear relation statable between the mastery of these specific materials and the larger aim of speaking Spanish.

III. Clear Statement of Assumptions and Principles

The assumptions indicate why the objectives have been considered desirable and feasible, and the principles show the logic underlying the specific procedures adopted in the course. Together the principles constitute a philosophy of language teaching.

IV. A Detailed Record of Procedures

Before a program can be the source of meaningful information, the procedure followed in it must be clear. Skeptics often claim that classes are different in name but not in activity. Certainly we have observed reading classes in which very little reading was done and oral classes in which no student spoke. The label is scarcely sufficient warrant for classification. Furthermore, if a group shows outstanding achievement, an investigation is primarily interested in discovering hypotheses which may explain its success. These hypotheses can be obtained only from a clear record of the training the class has had.

V. Accurate Measurement

An investigation assumes the existence of accurate tests and other devices for measuring the attainment of all important objectives. These must be valid and reliable instruments which give exact comparative data on the achievement of various objectives by contrasting procedures.

VI. Control of Non-linguistic Variables

Accurate measurement implies that non-linguistic variables are adequately controlled and that different results are attributable to the success or failure of the methods and materials used and do not rest primarily on other factors. Or, if these factors are closely related to language learning, their effect should be accurately measured.

VII. Hypotheses

As a result of the procedure outlined in the foregoing paragraphs, a series of hypotheses should emerge. These hypotheses could involve any of the points in this list. They might, for example, indicate that certain objectives were not obtainable, that certain principles were not sound, that certain procedures were not effective, and so on. Situations would then be found or set up which would allow these tentative hypotheses to be tested in still more tightly controlled situations so that further evidence of their validity could be obtained.

VIII. Conclusions

From the foregoing process certain conclusions (even if only of a tentative nature) should appear. These too might apply to any of the steps in the process. One could see how well certain objectives were being obtained, what procedures seemed more or less effective in producing given results, and the like. Scarcely less important than the actual conclusions would be the isolation of questions for further clarification, problems for further study, and hypotheses for further testing.

THE PRESENT INVESTIGATION

Our investigation was unable to meet all of these criteria perfectly. As we have indicated, the statements of the difficulties we encountered will be particularly valuable in indicating the need for further work. This list may also serve as a warning of points to be kept in mind when reading or hearing the many *ex cathedra* pronouncements commonly made concerning the success or failure of individual programs.

I. Coverage

We attempted, by letter and personal visit, to secure the co-operation of as many individuals and institutions as possible. We sought to get in touch with all language programs of which we learned, by either hearsay or formal announcement, but some courses may have eluded us. We made repeated public announcements of our testing program and furnished the tests free to any who were interested in using them. In some cases our repeated inquiries remained unanswered. Sometimes after rather prolonged consideration, institutions decided not to participate in the study. We cannot claim that our coverage is inclusive. We believe, however, that the stronger language programs were generally interested in demonstrating their success and in seeing how they compared with other efforts. We think that there were few, if any, very confident programs which failed to supply us with data. Even the students whose scores we took for comparative "norms" (who were not participating in experimental programs) were probably the products of the more interested and effective teachers.

We sought to cover as wide an area of languages and of scholastic levels as possible. Too often pronouncements have been based on experience obtained with a particular type of student, at a particular level of education, in a particular local situation, or under other specific circumstances. We wished if possible to collect available thought and evidence on the general principles which underlie all language teaching and learning. Since the colleges had been particularly stimulated by the work of ASTP, the bulk of our material concerns attempts of various institutions to adapt various features of this program to their regular instruction. The high schools, in general, seem to have awaited the results of the colleges' experience before undertaking similar experiments, and we secured relatively little evidence from them and from lower educational levels. For the upper levels of language instruction we were able to use the foreign students studying English in this country. Although many of them came with little background in English, most of them had two or more years instruction in their native country and hence represented roughly the level immediately above that of foreign-languages students here. We also worked with certain other groups wherever possible, but the bulk of our material involves these two classes of students.

We were also forced to limit the number of languages with which we worked. At the time the Investigation began, the possibilities of civilian instruction in Chinese, Japanese, and the other less commonly taught languages were still uncertain; hence, in order to secure numbers of students sufficiently large for meaningful results and to limit the time and labor devoted to test construction, we concentrated on English, French, German, Russian, and Spanish. Even for German and Russian we found it difficult to secure large groups.

II. Objectives

The objectives of many courses still lacked sufficiently precise definition. In large part, this vagueness was due to the fact that only recently have the aural-oral skills been generally emphasized. After many years of working with the reading objective, some consensus has been developed as to how reading ability can be specifically defined. The aims of speaking and aural comprehension show much greater diversity of interpretation and a greater variety in the materials used to achieve them. This fact caused us great difficulty in selecting programs and in choosing or constructing instruments which would accurately measure these important aims. The next two chapters on objectives and on the testing program will illuminate these problems in greater detail.

III. Statements of Principles

Actual classroom procedures frequently are followed because of habit or convenience, not because they have been judged sound means of achieving the objectives sought. Likewise, in many instances the objectives presupposed conditions which were not met by the school or by the student for whom the course was intended. It is a truism in the teaching profession that the best practitioners are often weak in theory, particularly in stating their own theory. We culled the literature and also asked some outstanding workers in the field to supply us with statements applying to their particular courses. We have clarified some issues here, but ambiguity at this point still makes it difficult to discuss many experimental programs.

IV. Procedures

In the case of some classes whose progress we followed most closely, members of our staff or observers hired by us recorded the classroom activities in detail, and an exact record is available. Yet this technique (like that of personal observation by the staff) was obviously too costly in both time and money to be used in more than a few instances, and we have usually had to be satisfied with much less precise information.

V. Measurement

As has already been mentioned, lack of clarity in regard to some of the objectives made precise measurement of them difficult or impossible. As usual, the task of selecting or developing a series of tests and other means of evaluation had the function of clarifying many of the linguistic issues involved in the experimental programs: the definition of speaking knowledge, the topics in which the student was to be fluent, the speed with which he was to speak and understand, and other major issues in teaching the aural-oral skills. For reasons which will be made clear later, it is now impossible to construct a set of aural or oral examinations which will be perfectly satisfactory to every one on all counts.

VI. Control of Non-Linguistic Factors

Data on many of these points is admittedly difficult to obtain under any circumstance: e.g., the effectiveness of the teacher (especially the reactions between his personality and the type of student found in that particular class). But even for relatively objective materials, precise data were often hard to obtain. Many teachers and students failed to give the information requested. Sometimes when the information was given, further checking showed that it was inaccurate. Even with the greatest possible care and trouble, the data available ran out before the theoretically desirable limit was reached. For example, some idea of the students' general intelligence was highly desirable. Yet the variety of tests used by different institutions for this purpose, the number of transfer students, and the many individuals who somehow failed to take tests presumably required of all students, produced many cases forwhich satisfactory data were not available. The size of this group will be astonishing to anyone who has never sought to track down this kind of information. In any case, perfect control of these factors would be possible only for an organization which could dictate (to fifty or sixty institutions) the full detail of their entrance applications, placement tests, and registrar's records. Similarly, it was impossible to control adequately the student's previous experience with the language. Often registrar's records did not cover incomplete courses or courses which had been failed. Likewise foreign residence or the opportunity to speak the language with family or friends could not be ascertained with the precision which was theoretically desirable. None the less, for each of the programs we covered, we present such information as we could obtain.

VII. Isolation and Testing of Hypotheses

The Investigation was not to set up programs of its own but was to follow experiments already in existence. Consequently, the general organization of these programs, the testing of them, and their other elements were all dictated by the needs and interests of the institutions and staffs which carried them on. They were not designed primarily to develop and demonstrate specific principles of language teaching. Probably most language teachers believe that, in educational experiment, education should have precedence over experiment. None the less, the study of experiments so organized and conducted must inevitably result in less clear hypotheses and less conclusive results than could have been obtained under other conditions.

VIII. Drawing Conclusions

A major obstacle to complete decisiveness of the findings is the lack of wholly adequate standards of comparison. In most instances non-experimental groups existing at institutions which also carried on experiments differed so markedly in type of objective, student, materials, and other respects that their effectiveness as "controls" is very slight. We sought to overcome this difficulty in part by securing a general standard of performance or set of norms, which will be described in Chapter IV. Though comparisons on this basis contribute important information, they do not demonstrate conclusively what would have been possible under different procedures but with clearly equivalent student body, teaching staff, and other facilities. Not only were these comparisons in cross-section lacking; longitudinal comparisons — those with earlier groups of students trained by other methods at that same institution — were likewise wanting. Few institutions had adequate records of this sort. Though some of them had regularly used the Co-operative or other reading tests, none had sufficient data on aural or oral achievement, either on such standardized measures as the Lundeburg-Tharp tests or on homemade batteries for which local standards had been developed. Some further causes of difficulty in obtaining conclusive results will be indicated in the following chapters.

The work of the Investigation demands two different, though closely related, presentations. The first treats primarily the various experimental programs, considered not only as individual undertakings but as a general movement in the field of language teaching. Since their achievement cannot be examined in a vacuum but must be viewed against the background of the general language situation, such a study must also consider the general level of attainment. The data derived from the experimental programs as well as from the more conventionally trained groups, the means used for obtaining this evidence, and the conclusions to be drawn from it are, consequently, the subjects of this first volume of our report.

But all this theorizing, experimenting, and testing has implications far beyond the particular experimental trends with which it was directly concerned. This activity has contributed to our general store of knowledge about the fundamental principles and problems of language teaching and learning. In addition to this evidence gathered from practical work in the classroom, the Investigation was also able to carry on some studies of specific problems under fairly well controlled conditions. These various findings too cannot be considered in isolation but must be fitted into our existing information. Consequently, the other volume of this report, SECOND-LANGUAGE LEARNING, summarizes our knowledge of language teaching and language learning as it exists in the present stage of this experimental epoch, enriched by the hypotheses and evidence produced in recent years.

Chapter II

The Objectives of Language Teaching

The objectives of language teaching are fundamental to any consideration of it, for discussions of procedures, texts, results, and all the rest can be meaningful only in relation to the ends sought. Objectives are particularly important for comparisons between the "newer" language programs and their predecessors because the fundamental difference between the two types of courses lies in their objectives.

Comparisons between the "old" and the "new" courses may, however, tend to imply that each type is more homogeneous than is actually the case. Not all language teachers accepted the findings and recommendations of the Coleman Report; and even those who sought to follow it produced very different courses. The "reading method," the "direct method," the "all-purpose course," and half-a-dozen others existed side by side, and very different programs were to be found under any one of these labels. Thus the "old" or "traditional" course was actually a motley array of rather different ones. The "newer" course is equally varied. The ASTP itself, though based on the army's general directive, was a collection of highly individualized adaptations and experiments as it actually existed at various colleges and universities; and the civilian programs developed from it have been no less diverse.

In spite of this diversity of both the "old" and the "new" courses, it is still possible to find in their objectives a major point of difference between the two general types. No scheme for classifying objectives is ever wholly satisfactory but for the moment the following distinction is useful. In one group of objectives may be classed the efforts to impart the basic linguistic skills: reading, writing, speaking, and aural comprehension. In the second group can be placed those larger objectives for which the language skills serve as means. A wide variety of outcomes has been listed here: disciplinary training in neatness, accuracy, and logical thought; increased understanding of the nature of language as a means of communication and as a tool of thought; a better command of one's native language; a knowledge of the foreign people's history, culture and civilization, especially a knowledge of the foreign literature; increased international understanding and good will; the development of historical and cultural perspective, and many others.

As far as this second group of objectives is concerned, comparisons between the "old" and "new" courses can have little solid foundation. Such comparisons have been made. For example, some critics of the "newer" courses have asserted that any superior achievement as regards the first set of objectives in these courses has been attained only at the expense of this second type. Yet there are no generally accepted standards by which to measure achievement in regard to this second class of objectives, and there seems likely to be none for some time. To take a typical example, "knowledge of the foreign culture and civilization" is an aim often stressed. Many teachers insist that they give considerable attention to it, and the publishers' blurbs for most texts make similar claims. Yet one has only to read the articles or listen to the speeches about this objective to realize that there are as many interpretations as there are authors and speakers and that language teachers are far from any agreement on what they intend by "culture and civilization." On the one hand, some mean "knowledge of the location of important cities, rivers, and districts," but others seem to have in mind a mystic empathy with the soul of the people. Discussions of "culture and civilization" tend, consequently, to break down in complete confusion because of either failure to define the terms at all or unwillingness to accept the definitions offered by others.

Under these circumstances our examination of language programs could not take into account this second class of objectives. This inability reflects a grave weakness in all language teaching and one which needs to be remedied. Since these objectives are the reasons for teaching the linguistic skills, language teachers should define them more precisely and be able to present better evidence that their students are attaining them. Meanwhile, comparisons between the two types of programs must rest solely on the different objectives chosen among the linguistic skills and on the various degrees of success achieved in imparting them.

Among the linguistic skills the newer courses, almost without exception, give greater emphasis to the oral-aural: the ability to speak the language and to understand it when spoken. All the other much-discussed features of these programs (informants, phonemic transcriptions, increased contact hours, smaller sections, and the rest) are simply means — once speaking ability has been adopted as a primary objective — of achieving this end. Provided the ability to speak is accepted as an important aim, many of these devices are obviously sound means to it. The contention will hardly be doubted, for example, that smaller sections, which give each student a greater individual share in the class's total time, will produce greater speaking ability. Other matters, like the phonemic transcription, will be subject to more debate. But the success or failure of the individual devices is important only in proportion to the emphasis placed on the objective which they serve, the ability to speak the language. The value of the oral-aural skills as primary objectives in language teaching thus becomes the basis of all further discussion — particularly, the relative emphasis to be given these skills in comparison with the other objectives of language teaching.

Possibly the best way of examining the contrast between the objectives of the older and newer courses is to begin by considering why the traditional courses did not stress oral and aural uses of language. These reasons can conveniently be divided into two classes. In the first are reasons why it was felt that the oral-aural skills *could not* be taught in regular educational programs even though they might be considered desirable. In the second group are the arguments for deciding that the oral-aural skills *should not* be taught as important objectives for language instruction. In opposition, the "newer" courses rest on the beliefs *(a)* that the problems or difficulties in teaching the oral-aural skills can be avoided or overcome, and *(b)* that certain changes in the situation now make oral-aural skills more important and desirable as educational objectives.

The first reason advanced for minimizing the oral-aural skills, even though they might be desirable and important, was the lack of time. The two-year language course was, and still is, that elected by most high school and college students. Five hours per week for a school year of 33 weeks probably constituted a maximum for most institutions; this schedule exposed the student to the language for 330 hours during the two years. Many institutions, with courses running only three hours a week and with shorter school years, fell to as few as 200 hours for the two-year period. This amount of contact was thought to be too short for imparting a speaking knowledge of the language.

Whether or not this belief was sound, the army program and the college courses developed from it have all provided more time. For ASTP, a minimum of fifteen hours per week for 36 weeks produced an impressive total. Though most adaptations have been unable to secure so large a block of time in civilian schedules, the newer programs are all marked by a *relative* increase in the total number of hours. In some cases this change has meant only an increase from three hours to five; in others, the number per week has risen to eight or ten, with very few courses getting more than ten. Whether this much additional time (still small as compared with that available in ASTP) is enough to develop oral-aural skills is another question and one to which later pages of this volume will contribute information.

A second cause for the traditional minimization of the oral-aural objectives was the customary size of language classes. To have thirty to forty students in a section was standard practice in nearly all high schools and was certainly not unknown in college classes. Yet if the entire class-hour were spent in speaking, each student would have a minute and a half at most — hardly enough practice for gaining any degree of skill. The newer programs, by using informants or drill-masters, have drastically reduced the size of sections to six to twelve students and in this way have made possible three to six times as much individual practice per class hour.

A third reason for avoiding the spoken language was the belief that many teachers, particularly at the lower educational levels, were insufficiently trained in it themselves. At the worst, the teacher's phonemic and intonation patterns were considered too poor to serve as models for students. At a slightly higher level, the teacher did not feel himself sufficiently fluent to be easy with oral-aural work. Yet there is a very real sense in which it is difficult for anyone who is not a native speaker of the language to handle an oral-aural course once it has progressed beyond the elementary stages. Subtle or idiomatic turns of expression present problems for which only native speakers of the language have feeling. By providing drill-masters who are native speakers, the newer courses have avoided this difficulty. The native speaker may not be competent to teach his language, but he is a judge of how things are or are not said in it.

A fourth cause for the earlier minimization of the oral-aural skills was the belief that the student had few opportunities to hear the foreign language outside the classroom. Here the newer courses have utilized the record, the movie and the radio as means (outside the classroom as well as in it) of enabling the student to hear some speaker other than the teacher and to come in contact with varied types of material. These courses have also done much to interest native speakers of the language (living in the community but not attached to the educational institution) in programs which permit the student to speak and hear the foreign language in social situations. Language houses and other means of this sort have been given renewed emphasis.

To sum up, in regard to all the reasons advanced for believing it impossible or very difficult to teach the oral-aural skills, the advocates of this approach have not so much denied these difficulties as they have sought to avoid or overcome them, and their innovations in procedure are best understood as specific means to these ends.

There then remains the second set of arguments: the earlier reasons for believing that the oral-aural skills, even if teachable, should not be primary objectives of language instruction. This argument is a question of relative, not absolute, value. All arguments favoring oral-aural work involve, either explicitly or implicitly, a comparison with reading. Probably no one thinks that ability to speak and understand a foreign language is a dangerous or undesirable thing in itself. At the high school and college level certainly, the possible psychological dangers of bilingualism or similar arguments against speaking a foreign language do not enter. The question has always been: are the oral-aural skills a *more* educationally desirable aim than reading ability?

The greater utility (educational or otherwise) claimed for reading rests on several points. One is that because reading is the easiest skill, it is the one which the student is most likely to acquire to a usable degree. In the attainment of any skill, a certain point of development must be reached before its possession is of much value. The student of piano, for example, who can play only *The Brownies' Picnic* finds that this much skill contributes little to himself or to others. His pianistic ability must reach what may be called a point of increasing returns. He can then at least make a stab at playing many of the things he wants to play and at accompanying singers or instrumentalists. The language skills resemble all other skills in that they too begin to produce returns only after a certain point of attainment has been reached; and past that point relatively small increases in skill produce greater returns in use and pleasure.

As we have said, this fact was used as an argument for reading because it is the easiest skill. Reading requires only the passive recognition, not the active manipulation, of the linguistic symbols. The reader can go at his own pace, pausing or glancing back if he desires, and he can stop to consult dictionaries and grammars. Being thus the easiest skill, reading seems to be that which the average pupil is most likely to develop to a useful point. No less important is the claim that reading is a skill which the student can practice by himself and that consequently it is not so necessarily subject to "rust" once formal training has ceased. Reading has also been favored on the grounds that it is the skill which the American student is most likely to find either the need or the opportunity to use for education, vocation, or pleasure.

The rebuttal offered by the advocates of the opposite position does not attack all these points. In general they agree that the aural-oral command is more difficult of attainment than is reading skill; but they believe that by devoting a larger portion of the student's time to language work and by exploiting mechanical aids and other new resources, they can overcome this handicap and develop the student's oral-aural ability to a point of increasing returns comparable to that he can attain in reading. Their denials concern chiefly the other two points and really rest on what they believe is a change in situation. Living in one world and enjoying modern technological developments in transportation and communication, the American student will no longer find (they claim) that reading is the skill which he has most opportunity to practice and use. They point to increasing vocational opportunities for which oral-aural competence is indispensable or desirable as industry seeks world markets and international co-operation in political and other spheres increases. They foresee greater social and recreational contact with other peoples as air transport and radio communication knit the world more closely together.

In addition to these claims of greater usefulness, they frequently urge two additional points. They believe that more students can be better motivated to study a language if that involves speaking it rather than reading it. Further, they claim that the initial approach to language, whatever the ultimate skill sought, is best made through the ear and tongue. Thus they believe that time spent in oral-aural work will increase reading ability within the same period of training.[1]

---

[1] The skills could stand in several different relations. For example, the aural and oral skills could be considered the dominant objectives, with reading relegated to a subordinate position. Though this relative emphasis was adopted by some military programs, only one or two extremely experimental plans for civilian instruction have adopted it. Most educational institutions continue to assume the primary importance of reading ability. Oral-aural skill is then regarded as either an introduction to reading or as an "extra

If one wishes to examine these counter-claims in their turn, it is obvious that certain reservations or qualifications must be made unless one is willing to adopt a completely partisan attitude. For example, vocational opportunities in language work have always existed, and there are probably more of them now than there were in the past. The question is whether they have so increased as to affect any large proportion of our high school and college populations and thus justify any great educational change based on these grounds. Special training programs such as language schools, intensive summer courses, and the like can always be available for the needs of those professionally concerned with languages. Needs of this sort must affect many students before they become cogent educational arguments.

Similarly arguments based on increased travel and communications sometimes represent future predictions rather than present actualities. Thanks to jet propulsion, air travel is rapid, but it is still so expensive that only a small segment of our population can indulge in it. Likewise, despite the five-day week, relatively few have vacations long enough to make extended travel possible. The same cautions can be urged in regard to radio communication. Short-wave is a fact, but how many people own a short-wave set? And not all present owners listen regularly — even to the BBC, for which no foreign language skill is needed.

---

dividend" which the student gets in addition to reading ability. Even in the first case, the exact relation which is thought to exist between the skills may be very different. Some see oral-aural skills as benefiting reading comprehension at even the most elementary levels; they believe that the student will read anything more easily and accurately if he possess "some" oral-aural skill. Others see the contribution of these skills to reading as coming much later — primarily when the student begins to read poetry and belletristic materials, and to read them for literary and stylistic qualities rather than for mere factual comprehension. Quite obviously it will make considerable difference which type of thing is meant when one talks about "aural-oral skills as aids to reading ability."

For deciding issues of this sort, factual data on the actual number of people concerned and other matters would be of inestimable help in arriving at sound judgments. The war and postwar periods, during which the Investigation operated, clearly could not furnish data which would be representative of the situation likely to prevail once the worst of the postwar dislocation is past. We did not, therefore, attempt to collect any such data. But if language teachers are going to use arguments of this sort, they should be ready to support them with evidence gathered as soon as general conditions make it possible.

The claim that greater emphasis on the spoken language produces better motivation in students likewise demands evidence as opposed to mere asseveration. There is a marked tendency for all of us to generalize our personal preference into a universal law. But motivation also suggests another important issue, the need for diversity. Some students may be more attracted to reading, others to speaking. In the best of all possible worlds, educational institutions would be able to present varied offerings which would meet the needs and interests of different types of students. Unfortunately, few institutions are able to offer a very wide assortment of courses.

The fourth argument of the advocates of oral-aural training also demands careful consideration. It asserts that the spoken language serves as the best introduction to any other type of language learning. This argument, however, hinges on a question of degree, and unless this degree is indicated exactly, battles with straw men are likely to result. Few traditional classes made *no* use of the spoken language.[2] Probably most teachers believe that the student, in order to read, must be able tomake some set of sounds to represent the words he sees. But reading with a consistent phonemic pattern is a long way from being able to carry on a conversation in the language like a native speaker. Likewise, some skill and practice in aural comprehension does not equal the ability to understand everything said. In the interval between the extremes, there are an infinite number of levels which teachers can consider the minimum of oral-aural proficiency necessary for the development of reading skills. To say that there should be an oral-aural base for reading is tó say little; the important question is how extensive (specifically) this base can and must be.

---

[2] For example, advocates of the "reading method" always stressed "oral exercises" though these were distinguished from speaking ability. Cf. A. Coleman, "A New Approach to Practice in Reading a Foreign Language," *Modern Language Journal* XV (1930) 101-18; P. Hagboldt, *Language Learning* (Chicago: University of Chicago Press, 1935) 44-5.

Both the layman and the professional language teacher sometimes speak glibly of "knowing" a language or "having a reading knowledge of a language" as if these abilities were unitary things which one either has or has not. Of course, in point of fact, knowledge of a language is a matter of degree, best represented as a line formed by a continuous series of points. At a given moment, any individual's knowledge of a language is represented by one of these points as he moves along this continuum toward the final point, "complete mastery of the language." Actually this last point is at infinity and is never reached, for no one ever attains "complete mastery of a language," even of his own. When anyone speaks of "a reading knowledge of a foreign language" or "ability to speak like a native," he obviously does not refer to the last point in this line, but rather to some other point or cluster of points much nearer the center.

Two difficulties are inherent in this situation. One is the problem of definition. When two people talk about "a reading knowledge," one of them often means some point quite far to the left or elementary end of the line while the other has in mind a point far to the right or advanced stage of language proficiency. Many past controversies in the literature have waxed hot as a result of this elementary semantic confusion.

But even after both parties have, so to speak, put their pencil-points on that part of the line they are talking about and after the referent of the term "reading knowledge" is clear, there is still plenty of room for disagreement about the second problem: whether one party has put his standards too high or too low. When is the degree of ability sufficient to merit the title of "reading knowledge"? Before turning to the newer problems involved in "speaking ability" and "aural comprehension," we do well to note the state of affairs in regard to reading knowledge. Since this objective has been familiar for many years, language teachers have had considerable experience in attempting to define it and to set standards for it. The knowledge and experience gained in regard to reading should be useful in doing similar work with the other skills.

Though any attempt to generalize must be open to numerous qualifications, common practice seems to define a "reading knowledge" of the Indo-European languages somewhat as follows.[3] The student can recognize approximately the first 2,000 words in the frequency counts made of written material in the languages for which these studies are available. In morphology and syntax, likewise, he is expected to have a "recognition knowledge" of the most common forms and grammatical patterns.

There then remains the question of judging the definition or standard so set: what does this "reading knowledge" mean when applied to the actual process of reading? Here the important point to notice is that the standard is conventional in that the student will be able to read very little without further resources and practice. Few authors except textbook writers will ever compose many consecutive sentences using only the vocabulary, morphology, and syntax of highest frequency.[4] Though the student will get some added help from cognates and parallels to his native language, from the sense of the context and similar aids, ultimately he will have to turn to the dictionary and grammar. Thus the "reading knowledge" at its early stages gives the student only the ability to slog through materials. With this much equipment he can begin to read in the language, and this experience will increase his familiarity with its vocabulary, grammar and syntax. If his reading centers primarily about a limited number of fields, this practice will ultimately give an approximation to a complete knowledge of the language commonly used by writers in those areas. Without dictionaries, grammars, and this additional practice, however, the "reading knowledge" amounts to little. As has often been pointed out, the graduate student who requires an hour to work through a page will never have the foreign literature of his subject at his fingertips. He must get additional practice and greater speed in reading before the "reading knowledge," as commonly defined and taught, means "being able to read." Because of the existence of the frequency counts and the relation between them and the usual definitions of "reading knowledge," this skill can be defined with fair precision. At the same time, the gap between the skill as defined and the skill in use should not be lost sight of.

---

[3] Cf. A. I. Frantz, "Reading Knowledge Test in the Foreign Languages: A Survey," *Modern Language Journal*, XXIII (1939) 440-6; M. S. Pargment, "What Constitutes a Reading Knowledge of a Foreign Language, and How it can be Achieved," *French Review*, XVII (1943) 74-82.

This general state of affairs in regard to reading (however approximate the foregoing statement of it is) is naturally the background from which one turns to consider the oral-aural skills. Much debated at the present time is the degree to which such transfer is possible. One frequently hears that the "spoken language" is not the "written language" and the degree of difference will, of course, be an important point in determining the amount of transfer which is possible.[5] Looking at the written and the spoken languages of Western Europe, one can easily see certain obvious points of difference. The spoken style is more loose than the written; sentences are shorter and simpler. Some differences also appear in morphology; for example, the past definite tense of French and the secondary subjunctive of German are certainly less frequent in the spoken language than in the written.

---

[4] M. West, "The Problem of 'Weaning' in Reading in a Foreign Language," *Modern Language Journal* XV (1931) 481-9.

The issue is sharpest, however, in regard to vocabulary. At certain points the frequency counts obviously reflect written materials rather than spoken ones. For example, the prevalence of knights, princes, and wolves may be largely due to the fables and fairy tales which frequently figure in the material used for the word-counts; and souls are perhaps more frequently written about than talked about.

As our investigation was being organized, the suggestion was frequently offered that actual counts, similar to the vocabulary and syntax counts existing for written materials, be undertaken of speech.[6] For various reasons which are not relevant here, these studies were never carried out; yet enough initial planning was done to reveal very clearly the problem of sampling involved.

---

[5] E.g., M. West, "The Present Position in Vocabulary Selection," *Modern Language Journal* XXI (1937) 433-7; F. C. Hayes, "What is a Basic Word?" *Hispania* XXII (1939) 307-10.

[6] Some word-lists for oral use, chosen primarily by a subjective alteration of the frequency lists for written materials, have been suggested in the past, but obviously they have had little influence on the texts and materials now being used in aural-oral instruction. Among the lists are: O. P. Schinner and H. G. Wendt, "A Suggested List of 1,000 Active German Words," *Germ. Quart.* VI (1933) 77-901; J. B. Tharp *et al.*, *A Basic French Vocabulary* (Washington: Modern Language Journal, 1934).

In making such counts, clear distinction would have to be made between three different types of conversation which could be recorded. If we take English as the language of illustration, then we could record English conversations *(a)* between foreign visitors and native English-speakers in the United States or the British Empire, *(b)* between foreigners and and native English-speakers in the foreigners' country, or *(c)* between native speakers of English.

The traveler seems primarily concerned with the problems of travel, shopping, and daily living. If we define "oral English" as that which the foreign visitor first needs when traveling in an English-speaking country, we will find it to be the language of the hotel, restaurant, and store. As a result, the phrase books and oral courses intended primarily for the traveler have always stressed "situations" or "topics" of this sort. Since the Army was training potential travelers in expeditionary forces, the ASTP materials also emphasized this area of oral communication, and this tourist talk is what many people have in mind when they refer to the "spoken language."[7]

In the second possible type of conversation, the foreigner would be speaking English to American or British visitors in his native country. In this case he would need the English in which to carry on his business or profession or in which to converse with them socially. He is not buying tickets or ordering meals. If he performs these services for his guests, he does it in his native language, not in English.

The third type of English conversation which could be recorded would be that between native speakers of English.

---

[7] Even the tourist, however, will find this much language adequate only as long as he has rather superficial contact with native speakers. If he is invited into their homes or otherwise encounters them socially, the talk is no longer about prices or vegetables but covers politics, the arts, personal gossip, and similar topics. The traveler's ability merely "to get around" will help him but little in these social conversations; but they are "spoken" too—as are movies, radio programs, sermons, and lectures.

If we listen to talk in a restaurant, we will find that the speakers order a meal, ask for the bill, and pay it. But between these little dialogues of arrival and departure, many minutes or hours intervene in which their conversation does not concern finding the railroad station or sending out the laundry. They are telling jokes, passing on the latest gossip, or doing any number of linguistic things other than using the phrases of the tourist. Readers can check this point by recalling their own speech for the past few days. Many of us have traveled, ordered meals, and asked directions. But even so, these activities have affected only a very minute part of our total flow of speech over a period of forty-eight or seventy-two hours. The bulk of our talk has been with family, friends, and fellow workers and has concerned an enormous variety of subjects.

If, therefore, we attempt to record "the spoken language," obviously the type of conversation we sample will directly control the kind of language we find. If we record speech at a ticket window, we will get a certain amount of the material usually embodied in lessons entitled "At the Railroad Station." But we will need to turn off our recording apparatus promptly. Otherwise the situational phrases will be buried in the flood of other conversation once the tickets are bought: discussions of the baseball season with the ticket seller, queries whether members of the family remembered to turn off the gas heater, and countless other topics discussed in the neighborhood of ticket windows.

Of course, whichever one of these situations is covered, there will be an almost complete overlap in the commonest function words: articles, prepositions, pronouns, and the like. But the content words (the nouns, verbs, adjectives, and adverbs) will vary considerably. Until extensive counts of oral materials have been made, predictions concerning them must be purely speculative. Yet the study made by the

Bell Telephone Company[8] and similar smaller counts for oral English suggest the likely outcome. To take but one example, some of the commonest English nouns, as determined by the telephone study (*order, price, car, letter, company, office, shipment*), clearly owe their rank to the fact that 80% of the conversations recorded were business calls. This result is not surprising since our existing literary counts tend to show the effects of sampling; a larger proportion of personal letters or newspaper articles, for instance, would have made considerable difference in the lower frequencies of the present counts of what are loosely called "written material."

If, in the collection of oral materials, sufficiently large samples were taken in order to rule out possible bias, an oral frequency list *might* emerge which would not differ very markedly in vocabulary and idiom from the ranks obtained for the written materials, provided the sampling for them too was increased to avoid what may be their present bias. Thus, in vocabulary and idiom, the "written" and "spoken" languages might not be so diverse as is sometimes claimed. As we have said, this hypothesis is pure speculation until adequate data for a judgment are available. We have advanced it here only to indicate how dubious are some of the assumptions frequently made about the differences between "the written" and "the spoken" languages.

Meanwhile, without exact frequency counts based on speech, a state of affairs has arisen in regard to the oral-aural skills which merits the closest attention. Most of the current oral-aural courses emphasize the situations in which the traveler finds himself; and their texts and other materials stress the vocabulary and idiom appropriate there. In part, they may do so because of a common tendency to confuse the

---

[8] N. R. French, C. W. Carter, Jr., and W. Koenig, Jr., *The Words and Sounds of Telephone Conversation*, New York: Bell Telephone Company Monographs, B-491, 1930.

"situational vocabulary" with the whole of the spoken language. (Yet many conversations about politics, daily life, or personalities are not too different in vocabulary from newspaper or magazine articles on these topics.) Pedagogical convenience may also favor the choice of situational materials. At the early stages of language learning particularly, classroom conversation (at best an artificial and highly perishable product) can best be nourished by "playing store" and using similar situations which lend themselves to dramatization. In the third place, teachers may feel that this language of "getting around" is really the phase of oral-aural skill most likely to be useful to students.

Unanimity does not exist even among those courses stressing the environmental vocabulary because they do not handle the same situations or handle them in the same fashion. Thus one Spanish class follows sport by going to the bullfight; another, by attending a soccer match. One class learns to order international items like beefsteak and French-fried potatoes at the restaurant, while another explores native dishes. These differences may appear petty but they make a very real difference in the kind of vocabulary the student controls at the end of two years. This is no new situation. About fifteen years ago, Tharp and his colleagues[9] found in their examination of seven conversational manuals that only 7% of the vocabulary items (161 words) occurred in all seven.

That the same diversity marks recent materials can be easily demonstrated. Three Spanish texts, intended primarily for oral-aural work, appeared in 1944 and were used by many experimental and conventional classes: Krakeur and Brugada's *Las Aventuras de Roberto Martín,* LaGrone's *Conversational Spanish,* and Treviño's *Spoken Spanish.*[10] In relation to the

---

[9] J. B. Tharp *et al., op cit.*

[10] L. G. Krakeur and I. Brugada, *Las Aventuras de Roberto Martín.* New York: Decca Records, Inc., 1944. G. G. LaGrone, *Conversational Spanish for Beginners.* New York: H. Holt and Co., 1944. S. N. Treviño, *Spoken Spanish.* Boston: D. C. Heath and Co., 1944.

Keniston List,[11] their vocabularies are distributed as follows:

DISTRIBUTION OF VOCABULARY OF THREE TEXTS FOR SPANISH AURAL-ORAL CLASSES

Keniston Classification	*Treviño* *No. of Words*	*LaGrone* *No. of Words*	*Krakeur* *No. of Words*
Group 1	499 (34%)	325 (50%)	290 (25%)
Group 2	245 (16%)	77 (12%)	146 (13%)
Group 3	177 (12%)	55 (9%)	133 (12%)
Group 4	179 (12%)	46 (7%)	135 (12%)
Total of Keniston's basic list of 2,000	1100 (74%)	503 (78%)	704 (62%)
Derivatives of Groups 1, 2, 3, and 4	100 (7%)	55 (9%)	122 (10%)
Total in Keniston's list of 3,060 words	1200 (81%)	558 (87%)	826 (72%)
Words not in Keniston's total list (3,060)	281 (19%)	87 (13%)	318 (28%)
Total Vocabulary of Text	1481 (100%)	645 (100%)	1144(100%)

---

[11] H. Keniston, *A Standard List of Spanish Words and Idioms*. Boston: D. C. Heath and Co., 1941.

All such counts and comparisons are subject to slight inaccuracies for which allowances must be made: proper names, compounds of slightly different form but similar or identical meaning, etc. Since we have followed a consistent plan, the biases of this sort are the same for all three texts. Nevertheless, the numbers are only approximate, though the percents are probably not affected by these matters, especially in regard to Spanish.

Since about 20% of the words in each of the three texts are drawn from the vocabulary beyond the 3,060 words of the Keniston list, it might seem as if the words chosen by the three authors would clearly suggest the additional (or different) vocabulary needed for aural-oral work. Such is hardly the case. If we consider the "non-Keniston" words as a group, we find that only eleven of them appear in all three texts: *autobús, avión, cerveza, deporte, equipo, frito, gasolina, gerente, patinar, resfriado,* and *tomate*. "Skating" is hardly a distinguishing characteristic of the spoken language, and the ubiquity of "tomatoes" seems equally fortuitous. The presence of "bus," "plane," and "gasoline" probably reflects only developments in transportation which would also affect modern written materials. So the list of things more talked about than written about (such as "colds," "managers," "sports," "teams," etc.) which are suggested by this unanimous selection is not long.

Though the small vocabulary of LaGrone's text may seem to give a spuriously small figure for the degree of overlap in the "non-Keniston" words of the three vocabularies, the results can be substantiated by further comparisons between any two texts. For example, LaGrone uses approximately 87 of these words, but only 19, or 22% of them, also appear among the 281 words of this type used by Treviño — only 8 more of these words than the 11 which appear in all three texts. Similarly, both Treviño and Krakeur use a fairly large number of these words, 281 and 318 respectively; but only 46 of them (16% of those used by Treviño and 13% of those used by Krakeur) are common to both.

Similar studies of the texts for other languages reveal the same situation. The texts developed for aural-oral work do not follow the frequency lists, yet they do not agree on the vocabulary content of the elementary "spoken" language. The cause of this disagreement is clear. Most "situations" actually demand a full command of vocabulary. The housewife "at the market" must know the scores of words denoting common fruits and vegetables; but the author of a single elementary text can select only a dozen such words. It is hardly surprising that these choices are not unanimous.

Meanwhile classes working primarily with reading materials and having their aural-oral drills based primarily upon them will be familiar with a vocabulary much closer to that of the written counts. The possibility of covering both types of vocabulary (the environmental and the frequency) exists, and most programs which use a prolonged oral-aural introduction but retain reading ability as their primary aim, rest on the hypothesis that this combination can be achieved. Yet if the "written" and the "spoken" vocabularies are as distinct as is usually assumed, the learning burden of elementary vocabulary is doubled, and courses covering both areas will do less with either than courses which give primary emphasis to one.

One possible solution to some of these difficulties is the "minimum adequate vocabulary" or "vocabulary of maximum coverage," a device used by many workers in the field of English. Possibly the most familiar of these is Basic English, which by general terms, circumlocutions, and definitions claims to be able to express all concepts through a very limited number of words. Little work has been done along similar lines in foreign languages. These small-but-mighty vocabularies have obvious advantages in controlling reading materials and in enabling the student to express himself in speech and writing since for these latter skills he can stay within the limits of the system which he knows. But their lack of popularity in the foreign-language field may be due to the same objection urged against them in English: their shortcomings in regard to aural comprehension (for the student can understand only a speaker who knows and keeps within the system the student knows) and their supposed clumsiness and lack of colloquial tone in oral expression.

The issues can be exemplified by taking the definition of one worker in the field,[12] who is particularly interested in the oral-aural approach:

> A person has "learned" a foreign language when he has thus first, within a limited vocabulary, mastered the sound system (that is, when he can understand the stream of speech and achieve an understandable production of it) and has, second, made the structural devices (that is, the basic arrangement of utterances) matters of automatic habit. This degree of mastery of a foreign language can be achieved by most adults, by means of a scientific approach with satisfactorily selected and organized materials, within approximately three months. In that brief time the learning adult will not become a fluent speaker for all occasions but he can have laid a good foundation upon which to build, and the extension of his control of content vocabulary will then come rapidly and with increasing ease. [Underlining in the original.]

The scope and content of this limited vocabulary will determine the kind of "mastery" thus intended, for the specific selection made will influence not only the content words, but will also dictate (to some degree) the structural patterns the student practices. Most important, the student who has "learned" a language as thus defined, can certainly not talk about everything or understand what is said to him — except within the particular vocabulary chosen. "Understanding the stream of speech" also raises questions regarding the speed, clarity, and amount of the speech heard.

In short, as regards the definitions of the oral-aural skills, we have yet to arrive at as clear a definition as in the case of reading (despite the obvious shortcomings and difficulties of that definition). The vocabularies and constructions chosen do not have a common reference point as those of reading have in the frequency counts. And those which have been selected by individual teachers or staffs have been based on differing principles.

---

[12] C. C. Fries, *Teaching and Learning English as a Foreign Language*. Ann Arbor: University of Michigan Press, 1945, p. 3.

The problem of judging the adequacy of such a common standard or definition, however, assuming that it could be arrived at, presents one special difficulty in the oral-aural area. As we have seen, the "reading knowledge" is expected to be open to considerable supplement; the student will be able to look up words, forms and constructions unfamiliar to him. In speaking and listening, such pauses for reference are usually not possible. Though some conversations have been carried on with much thumbing of the dictionary, usually the speaker will not wait while the hearer looks up words; the latter must get the language on the fly. Much the same is true of speaking. The speaker of a foreign language cannot take hours to utter a sentence, especially in view of the fact that he is usually better understood if he manages to rush his words together in the way the native is used to hearing them. The listening and speaking standards, therefore, must be *higher* than that taken for "reading knowledge" because they cannot be supplemented at need on the spot.

Thus in the oral-aural skills there is often a wider gap, for example, between "aural comprehension" and being able to understand any normal speech of a native speaker than there is between "a reading knowledge" and being able to read any given material in the language. This problem vexes all discussion of the newer courses and frequently gives rise to complete misunderstanding. Such confusion would certainly result if an outsider expected that after three months the student who had "learned English" according to the definition just quoted would be able to say anything he needed to say or be able to understand almost anything said to him.

To sum up, as the reader examines the "experimental" and "conventional" courses reported in the following pages, he will find that the former give much greater emphasis to intensive instruction in the aural-oral skills. But these courses will differ considerably in their specific practices in regard to

the content of the "spoken" language, the degree to which these skills are to be attained in two college years, and the purposes which these abilities are thought to serve. Needless to say, unless the next few years bring greater clarity on all these points, the present confusion and contradiction will continue.

Chapter III

The Measurement of Language Proficiency

Within the competitive system which characterizes American culture, it is not unnatural for teachers, along with other producers of goods and services, to proclaim the excellence of their product. Lately, teachers of modern foreign languages, experimenting widely with new methods and techniques, have not hesitated to claim superior results. These claims have tended, however, to be based on teachers' subjective judgments or, if supported by concrete data, on evidence drawn from tests devised by the teachers themselves in connection with the particular materials used in their own courses. If the Investigation was to appraise objectively the attainments of students in a wide range of experimental and conventional courses, it needed more than personal claims and estimates. We ourselves were able to attend the classes of several experimental programs; but, however impartial we might hope to be, our personal standards would be little more valid and reliable than those of individual teachers. In order to judge the success of instruction, we needed as accurate, impartial, and comprehensive a testing program as could be devised.

The majority of the newer courses which we sought to evaluate set as their initial objective the ability to handle the spoken form of the language — that is, the development of the aural and oral skills. In few, if any, of these programs, however, were the aural-oral skills regarded as the final or only goal. The ability to read widely in the language was usually retained as the most desirable ultimate objective. Thus in both experimental and conventional courses the measurement of the reading skills was a common ground for testing.

READING TESTS

As the result of development over the past decade and more, there exists a fairly well established level of achievement as regards the difficulty of content and the range of vocabulary and grammatical construction to be mastered for reading purposes. For the commonly taught languages, numerous textbooks have been prepared in accord with the word-counts, idiom-lists, and syntax-lists, which form the basis for this standard, and objective tests have been developed along these same lines. Among these are the Co-operative Tests published by the American Council on Education, which have been in nation-wide use for a number of years, with norms established for each of their parts: reading comprehension, recognition vocabulary, and functional grammar. To measure these skills on the basis of the one-year and two-year norms for these tests seemed to us a satisfactory method of appraisal which would eliminate the need for devising and standardizing similar tests of our own.

The number of college students who pursue the study of languages decreases sharply at the end of one, and especially two, years because the majority of colleges require only these amounts for graduation. Since, consequently, the end of the first and the second years were the points at which the most testing data would be available, we decided to administer the advanced level[1] of the Co-operative Tests at both these points. Scores are relatively low on the advanced forms at the end of the first year; but since satisfactory norms are available at that stage, we preferred the advantage of comparable scores for both years, as compared with the alternative of using the easier elementary forms the first year and shifting to the advanced forms at the end of the second.

---

[1] Revised Series, 1937. In French, Advanced Forms N, O, P, Q, or R; in German and Spanish, Advanced Forms N, O, P, or Q.

Since our studies were also to cover Russian, we developed a Russian test on the model of the Co-operative series, which at that time did not include that language. No frequency count of written Russian existed when this test was constructed; the author,[2] an experienced teacher of college Russian, had to rely on her own judgment in setting a level of difficulty which would be, part by part and item by item, generally comparable to that of the Advanced Co-operative Tests in the other languages. The Russian test was administered to students in seven colleges, and tentative norms based on this sample are included in Appendix A along with the norms for the Investigation's aural tests.

TESTS OF AURAL COMPREHENSION

Up to the present time, standardized measures have been in general use only for testing the reading skills. Even some of the early experimental courses with aural-oral emphasis were compared with more traditional programs solely on the basis of the students' ability to handle the written language. This limited testing was inadequate to our purposes. Students who have been taught to control the spoken language, even though they may ultimately demonstrate as high a reading proficiency as students trained primarily for reading, are assumed to possess additional skills which the other students lack and which it would be desirable to appraise, especially if it is claimed that learning to speak and understand the language is the chief source of those students' motivation. Furthermore, we did not feel justified in assuming so high a degree of correlation in individual students between reading and aural-oral skills that a high score on a reading test would necessarily indicate equal ability to speak and understand, or vice versa. As a result, we felt the need for a series of tests which would measure directly and reliably the aural and oral abilities.

---

[2] Frances Marshak-Sobotka.

In the aural field, the pioneering efforts were the Lundeberg-Tharp Audition Tests for French, German, and Spanish. Designed to measure aural skill on an elementary level, these tests had been administered widely enough throughout the country to have an established set of norms for the end of the first and of the second year.[3] Despite the existence of these norms, we questioned their objectivity because of the varying ways in which the tests had necessarily been administered. The material to be comprehended is spoken by the teacher giving the test, and thus there is little assurance that students in different schools hear the material delivered with comparable accents (native or non-native), at comparable speeds, and with comparable facial and other gestures.

During the war the testing staff of the US Armed Forces Institute had begun a set of aural tests modeled on the Lundeberg-Tharp series. This organization was disbanded, however, before the tests were completed although three preliminary forms had been prepared in several languages and had been tried out in various high schools and colleges. Though we made extensive changes in the form and content, we managed to avoid some duplication of effort by being able to secure such preliminary tests and data as were available at the close of that project. But these tests too had been administered by local teachers, and there was some evidence that considerable variation in student achievement could be attributed directly to the different ways in which the materials had been read.

---

[3] For descriptions of these tests, see R. D. Cole and J. B. Tharp, *Modern Foreign Languages and their Teaching* (New York: D. Appleton Century Co., 1937) 342ff.

In contrast to recent reading tests, which are based on a well-defined series of criteria as regards vocabulary, idiom, and syntax, the aural tests which we have mentioned could be based on no similar canon of materials common to the aural-oral method because none had been established. Pioneers in the construction of aural tests had had no criteria other than the counts based on samplings of the written language. These items, though likely to appear frequently in reading texts, are not necessarily encountered in any single book designed for the conversational or aural-oral approach. These textbooks themselves were comparatively few in number; and while each endeavored to embody the kind of words and patterns most commonly used in speech situations, their respective authors had quite differing notions as to what makes up the most important vocabulary of speech.[4] We too had no choice but to adopt the criteria used by the earlier tests, the existing lists.

The most important contribution we could make within the time permitted us was to standardize the spoken materials by recording them on phonograph discs, thus making certain that students everywhere would be hearing the same voice at the same speed, and with the same absence of nonlinguistic clues. This technique, though greatly improving standardization, certainly does not render it perfect. We sought to overcome some of the difficulties. The recordings themselves were the most faithful that could be produced, made in a well-known commercial studio under the direction of technical experts. Sets of instructions were issued to administrators of the tests, specifying what conditions should be met, such as the use of a high-quality electric phonograph with a light tone-arm, a metallic shadowgraphed needle, a room with satisfactory acoustics, with the loudspeaker placed at ear level in front of the class. The fact remains, unfortunately, that these instructions were not always followed; and in any case, not all rooms and phonographs are equally good. The conditions under which the tests are actually given ranged from laboratories carefully prepared by acoustical experts to the most casual of arrangements in which the records were played on ancient portable phonographs in run-of-the-mill classrooms. Airplanes, riveters, and sound-trucks sometimes added complications. It is, of course, impossible to calculate the differences in results produced by these variable physical conditions, and we have had to trust that they did not constitute a major factor in the bulk of the scores.

---

[4] Cf. pp. 32ff. above.

Form and Content

Two tests were constructed in each of four languages: French, German, Russian, and Spanish. They consist of an elementary, or "lower level," and an advanced, or "upper level," test.[5] Each test is composed of three parts, and may be administered to a group of any size in a maximum of fifty minutes.

The lower-level tests were designed for students who had accumulated a total of between 90 and 130 class hours in college. Part One consists of twenty-five completion items, each with a three-choice response (worded in English); for example, the voice says: "*Se cultivan las flores en un....,*" and the student answers by choosing among the expressions *garden, box-car, coal-mine,* printed in the test-booklet. Part Two consists of twenty-five definition items, each with a three-choice response; for example, the voice says: "*Ein Mann, der Fleisch verkauft,*" and the student chooses his answer from among *baker, butcher, doctor*. Part Three is made up of six short anecdotes, with from five to nine triple-choice responses, worded in English, on the content of each.[6]

---

[5] In French and Spanish, both the lower and the upper level tests were made in two forms (A and B) designed to be of equivalent difficulty.

The upper-level tests were intended for students who had received more than 150 class hours of instruction. Part One consists of twenty-five definitions; Part Two is composed of six anecdotes with from six to nine triple-choice responses on each; Part Three consists of a five-minute dialogue between a man and a woman speaker, with fifteen triple-choice responses on its content.

Problems of Construction

We found ourselves confronted, as we developed these tests, with a thorough lack of accepted standards as regards (1) how fast a flow of speech, (2) how long an utterance, (3) how wide a range of spoken words the student can be expected to understand at specified stages in his learning process. Earlier aural tests had failed to set such standards, and without them we had no accurate picture of the range of aural ability our tests would be expected to measure. Our work was to be primarily with intensive courses, providing anywhere from six to ten hours of conversational practice with native speakers, which were still in the blueprint stage. The ASTP experience had raised high the hopes of oral-aural enthusiasts among the language teachers, who foresaw their charges attaining near-native expertness at the end of one, or at the most two, concentrated years. Thus, although we had no actual basis for knowing what the level of aural performance would be, we could hardly ignore the widely made assumption that in many specially taught groups it would be high. For these students our tests would need a high ceiling. At the same time our examinations would be taken by large numbers of students conventionally taught without emphasis on aural attainment. Suspecting that in any case the range of skill to be measured would be considerable and doubting that a single test could cover it adequately, we constructed the tests at two levels in each language.

---

[6] The make-up of the French Lower-Level test differs from that of the other languages in that the Completion Series is replaced by thirty Phonetic Accuracy items, each with a four-choice response; for example, the voice says *parlent* and the student chooses his answer among the forms *parlons, parlâmes, parlent, parlant*, appearing in the test booklet.

In establishing standards of speed and length of utterance, we endeavored to base our criteria on normal speech. We argued, in the first place, that speakers of a language (other than elementary teachers of it) rarely distort the rhythm of their speech by introducing superfluous or unnatural pauses within normal word-groups. We therefore made sure that such choppy speech would not be heard in our aural tests by directing the (native) speaker to keep his natural habits of rhythm and juncture, and to reduce only slightly his normal conversational speed.

We further argued that often a person is called upon to apprehend speech sustained through considerably more than a single sentence — for example, when listening to a lecture, radio broadcast, or sound-movie. For this reason we included the anecdotes, which are two or three minutes long, as well as (in the upper-level forms) a five-minute dialogue.

In the selection of vocabulary, we were forced, as we have already noted, to rely on the existing word-counts[7] since no other criteria were available. There is far too little overlap of vocabulary from one to another of the conversational textbooks now in use to provide a common body of words of any appreciable size. In the lower-level test we restricted the vocabulary to the 1500 most frequent words, though adding a few items admittedly commoner in spoken than in written idiom, and a few cognates of English words. In the upper-level tests the limit set was the first 3000 words. However, the use of the high-frequency ranges of the word-counts resulted in a very small incidence of such vocabulary items as form the core of the "spoken language" textbooks. The vocabulary of these texts is largely of the type employed when a person is ordering a meal, sitting in a barber's chair, or trying to get his car fixed, and it consequently tends to omit many of the general words which have been found commonest in the written language. Thus, for example, of the total word-content of the Spanish Lower-Level test (Form A) only 36 per cent is found in the first half of Treviño's *Spoken Spanish;* of the total word-content of the upper-level test (Form A), only 77 per cent occurs in the entire book.[8] Despite this situation, it was evident that if our aural tests were to be used generally — that is, by students taught by many different procedures and materials — it was wholly out of the question to introduce carburetors and cauliflowers into selections designed to test aural comprehension. On such a basis, no matter what the particular choice of specialized words, the tests would necessarily have proved unfair not only to students of high-frequency vocabulary, but also to users of most aural-oral textbooks other than the one happening to embody the particular words occurring in the tests.

---

[7] For French: Vander Beke, *French Word Book* (New York: Macmillan, 1941). For German: Morgan, *German Frequency Word Book* (New York: Macmillan, 1937) and Wadepuhl and Morgan, *Minimum Standard German Vocabulary* (New York: F. S. Crofts, 1944). For Spanish: Buchanan, *A Graded Spanish Word Book* (Toronto: University of Toronto Press, 1927) and Keniston, *A Standard List of Spanish Words and Idioms* (Boston: D. C. Heath, 1941).

In addition to actual auditory apprehension of a spoken utterance and knowledge of the meaning attached to the sounds composing it, adequate comprehension very often depends on the listener's awareness of certain facts or assumptions bearing on the subject of discourse but not uttered in the actual speech situation. In the testing situation, however, it seemed to us imperative to eliminate this variable. Consequently, in order to respond to the test items, the only knowledge required, beyond what is contained in the actual utterance, is that which is probably common to all adult members of American society.

---

[8] For Form B, the corresponding amounts of overlap are: Lower Level, 43 per cent; Upper Level, 60 per cent.

Obviously many types of response made to comprehension in real situations cannot be used in an objective test. Oral responses cannot be made to serve as a measure of comprehension, since adequate answers may themselves be inhibited by numerous non-aural factors and since these answers, however appropriate, cannot be scored objectively. Consequently substitute responses must be utilized in the form of multiple choice. This substitute type of response, while eliminating the variables inherent in the spoken responses, must in turn be kept free of obstacles. One such obstacle would lie in the written form of the language being tested, especially since many programs planned to postpone or minimize work with normal orthography. To avoid this hazard, all the answer-choices of the tests appear in English, thus assuring that failure to answer a comprehension item is not due to inability to interpret correctly the choices themselves. The bilingual procedure involved here is seemingly not a hazard; most first and second year students, even while listening to a foreign language, are doubtless thinking in English and will not find their mental processes wrenched by the sight of English answer-choices.

Every attempt was made to reduce to a minimum other difficulties which the students might experience. Thus they were first allowed to hear a warm-up passage, which was not part of the test proper, in the hope that it would accustom them to the strange recorded voice and help them relax their initial tensions. In the body of the tests, all part-numbers, item-numbers, references to exercise-groups, and the like were given by the speaker in English, so as to help the student keep his bearings.

Reliability

The reliability coefficient is an index of the degree to which the score of a student on a particular test may be taken as an accurate prediction of his scores on a series of equivalent tests. Unless all these scores were similar, any one of them would be an inaccurate measure of the ability covered by the tests, say, the ability to understand spoken French. This coefficient is usually found by calculating the correlation between the scores made by a group of students on two equivalent forms of the test; or, when only one form is available, it is divided into two parts on the basis of the odd versus the even items or by some other method of division which will produce equivalent halves. A number of complications affect this procedure in any kind of testing.[9] Four factors merit special notice here because they are connected with the problems of aural testing and because they tend to lower the coefficients reported in Appendix C.

Aural testing produces a somewhat unusual situation. In taking most tests, the student first works rapidly through a section, answering those items about which he is certain and returning later for a more careful consideration of those items about which he feels less sure. In an aural test this procedure is impossible. The student cannot turn the teacher or the phonograph record back to the particular items he may want to reconsider. He must deal with each one at the time it is spoken; and, though considerable time is given him in which to record his answer, he must eventually choose something or else leave the item unanswered. It is practically certain that his auditory memory will be unequal to the task of allowing him to go back at the end of a part or of the entire test in order to work on items he has omitted. This inexorable progress of the test may lead the student to guess. Though judicious guessing improves the student's score and often produces a truer estimate of his peripheral knowledge than does a too cautious approach, the aural situation calls forth wild rather than judicious guessing because the student feels he should take a flyer while he has a chance at the item. The customary formula used to correct objective test scores for guessing[10] serves, on the average, to adjust the total scores, but the part-scores obtained by the split-half technique are much less accurately adjusted. As a result, tests in which guessing figures will appear even more unreliable than the guessing actually makes them.

---

[9] Cf. L. J. Cronback, "A Case Study of the Split-Half Reliability Coefficient," *Journal of Educational Psychology* XXXVII (1936) 473-80.

This factor is closely connected with another which we have already mentioned several times: the vocabulary and other linguistic elements to which students have been exposed vary widely from course to course and text to text. As a result, this fact that many words and constructions in each test were unfamiliar to the members of any particular class probably further increased this tendency toward guessing.

A third complication in constructing tests of maximum reliability was that the examinations had to be built prior to the experimental programs which they were to evaluate. Because many teachers were hopeful that new procedures would enable the aural-oral groups to far surpass students' previous attainment in all the skills, the tests were made more difficult than would have been desirable in view of the performance of those conventionally trained students on whom they were tried out in the preliminary forms. Unfortunately, as we shall see in later chapters, the aural mastery of the experimental groups was frequently not markedly superior to that of the conventionally trained students. Hence, providing scope for the anticipated higher level of performance served only to make the tests too hard, and consequently somewhat less reliable, especially for the conventionally trained students from whose scores the coefficients were computed.

---

[10] The formula is: Corrected score $= \text{Right answers} - \frac{\text{wrong answers}}{\text{choices offered} - 1}$; that is, a test in which three multiple-choices are offered is scored — rights minus one-half the number of wrongs.

The fourth feature is that correlation coefficients are sensitive to extreme deviations. The figures we report, however, are based on groups of students rather homogeneous in ability and in amount of training. Had the coefficients been computed on more diverse groups, the figures would be higher — spuriously so, however, though this procedure is sometimes followed.

In sum, the reliabilities reported for the tests tend to be minimum estimates and are higher for particular classes in situations where some of the foregoing factors have been minimized. Certainly more reliable measurement is imperative. We would only point out, however, that in view of the chaotic state of theory and practice in regard to the aural and oral skills, the coefficients obtained are almost surprisingly large. Since scores were reported for the individual parts as well as for the total tests, we have given their reliabilities also. In most instances, however, teachers used only the total scores, which are, of course, the most reliable.

Evaluation of the Aural Tests

Having discussed to this extent the problems of constructing aural tests, we hardly need to disclaim perfection. Aware that the tests were necessarily of a tentative character, we solicited criticism from their users, and it was given without stint. Since the nature of the tests conditions the

2

results obtained from them, it is highly important to examine this criticism and to attempt to distinguish the sound elements of it from what we consider the more questionable objections.

Some of the completion and definition items, it was pointed out, depend for correct response on the apprehension of a single word. For example, on hearing the item "*Celui que l'on paie pour qu'il se taise,*" one cannot choose between *lawyer, preacher,* and *blackmailer* (the three answer-choices) unless one catches the final word, *taise*. In spite of our efforts to avoid them, a few such items appear in one or another of the tests. The item just cited is also characterized by a certain trickiness which is best avoided.

The six anecdotes in each test were, for the most part, lifted from rather antique published collections, and many of them qualify as "old chestnuts." They tend to deal with the amusing quirks of imperious monarchs or eccentric men of letters, or to illustrate the wittiness of eighteenth-century courtiers. This is subject matter of a kind that the student has little familiarity with or interest in. If he happens to have read the anecdote in his language text, he has an obviously unfair advantage over his fellows. If not, he is baffled by the bookish flavor of the prose, which in some instances is strictly eighteenth-century, full of literary turns of phrase (especially the German) or of unfamiliar verb forms (e.g., the French preterite). In the Spanish Upper-Level test, some of the anecdotes appeared in *Selecciones del Reader's Digest* and were thus familiar to many students. Some of those in the Russian lower-level test have their source in a standard grammar used by many, though by no means all, learners of Russian. And while all these anecdotes were edited so as not to exceed the stipulated ranges of vocabulary-frequency, they still contained many words (e.g., French *coquin*) rarely heard in modern speech. Another factor is the essential structure of an anecdote, the whole significance of which is usually contained in its last sentence. This leads to considerable difficulty in phrasing multiple-choice statements on its content in such a way as to avoid giving away the main point without unfairly disguising it. Students who, in the actual listening, have wholly missed the "punch" of an anecdote will often be fairly successful in guessing the correct answers from clues in the wording of the multiple-choice statements. These various considerations have led us to favor instead a more realistic "slice of speech," consisting of dialogue or expository prose.

The five-minute dialogue constituting Part III of the upper-level tests represents one such type of selection on a larger scale. These dialogues have been criticized as entailing, for correct response, too much mere memory of detail to constitute a valid test of aural comprehension. However, the passages were carefully selected as having logical cohesiveness and unity; the answer-statements concentrate on the outline but not on the minute details of the conversation, and inquire about the characters and attitudes of the speakers in the situation; thus if the student has truly understood the dialogue, it is not a feat of memory to work out the answers. They are in his consciousness as an inevitable consequence of real comprehension.

A defect of a different order is to be found in the passage at the beginning of the first record in each test. In this passage the speaker announces that in order to allow the listeners to familiarize themselves with his voice and manner of speaking before the test begins, he will read the directions for Part I in the foreign language. This "warm-up" passage invariably had an effect exactly opposite to that desired. Instead of relaxing and accustoming themselves to the speaker's voice, the students were more unnerved than before because they understood nothing of what was being said.

This was of course merely because the vocabulary of directions was unfamiliar; but the psychological effect was bad as the students approached the first test item with the fearful conviction that they would continue to comprehend nothing. There is evidence that many did not overcome their nervousness until the first part of the test was over and tension was released between that and the second part. It seems clear now that the warm-up passage should have consisted rather of a series of easy sample items of the same type as those making up the body of the test.

Criticism which seemed to us less justified was aimed at the level of difficulty at which the tests were set. A frequent complaint was that the speech on the records was too fast for the students to understand. Despite this protest, the speech is of normal conversational rapidity in the upper-level test, and somewhat slower in the lower-level. Many teachers argued that it was not justifiable to insist on a normal speaking rate when, as in the case of records, the ordinary visual aids to comprehension — the speaker's facial expressions, his gestures, much of his personality — have been artificially removed. To this we may answer, first, that radio speakers seldom feel obliged to reduce the speed of their delivery because they are unseen by their audience; and, second, that a person listening to a high-fidelty recording or broadcast in his native language rarely complains of difficulty in understanding the unseen speaker. What the protesting teacher is really contending, then, is that his students can understand the language only under limited conditions. When he tells us, for example: "I suppose that the phonograph records are largely designed to take the place of native speakers; they would, therefore, scarcely be necessary in our case," he is at least partially motivated by the thought that, native speaker though he is, his charges would perform far more creditably under his familiar and studied (perhaps in this case even *over*-studied) delivery than when confronted with the unfamiliar and unembellished speech on the records. There is little doubt that they would, but this seems immaterial. Our object was not to test the adeptness of a given class at understanding their own teacher displaying his familiar tricks; it was rather to measure with a common yardstick the aural ability of language students from Maine to California. Such a common yardstick required a single voice speaking at a standardized speed. Through its use, we could at least determine objectively the extent to which, under the conditions it imposed, students of a given language understood that language in its spoken form. If "aural comprehension" means only "the ability to understand a familiar voice, speaking very slowly and overenunciating," this limitation demands more recognition than it has received.

In summary, most language teachers have tended to think of the aural tests more or less directly in relation to their own particular course, textual materials, and students. These courses and materials, if not the students, are so diversified that the tests were bound to impress the individual teacher as unfair to his students. But he must be reminded that if his own full battery of suggestions for revising the tests were to be adopted, his colleagues directing other courses would then find them even less satisfactory than before. Given this existing diversity, we could ourselves only hope to devise tests equally unfair, if not equally fair, to all potential users.

TESTS OF ORAL PRODUCTION

In the field of speaking ability, we were faced with even fewer established standards of achievement than existed for aural comprehension. There was no consensus among teachers regarding the range of vocabulary or the degree of grammatical complexity which the language student could be expected to use actively at the end of the first and the second year. As for actual tests of oral production, we knew of none in published form for general use. A group of testing experts at Queens College had been commissioned to construct a series of oral tests for use in the Army Language Program, and while these tests were never fully developed, it seemed desirable to draw upon the experience of the Queens experts and to have them prepare the oral tests.[11]

The tests were designed to measure speaking ability at the same two levels as the aural comprehension series: (1) the lower level, intended for students who had had less than 13[illegible] hours of classroom instruction: (2) the upper level, designed for those who had accumulated more than 150 class hours.

As in the case of the aural comprehension tests, and for the same reasons, there was no fair basis for vocabulary selection other than the standard word-lists. Working from these, we used the same criteria as for aural comprehension, i.e., up to 1500 words on the lower level and up to 3000 on the upper level. Difficulties of syntax were not intended to exceed those of vocabulary at the respective levels.

Each test was designed in three parts, intended to measure three essential aspects of oral skill: (1) the ability to report a single, simple act or situation in precise words, (2) the ability to express a sequence of ideas fluently, (3) the ability to converse. As will be noted in the rating-scales reported below, the major criterion in all three aspects was intelligibility.

Part I, the Picture Series, consists of twenty separate pictures, each representing a single simple action which can be reported in a single simple sentence. Two sample pictures, shown at the beginning, provide below them printed answers which would be normal if expected in English, e.g.: "The man is waiting for the train," or "The mouse is eating the cheese." The student makes his response to one picture before going to the next. His answer is rated by the examiner, as objectively as possible, according to the following scale:

---

[11] This group was composed of H. N. Rivlin, I. A. Schwartz, H. Schueler, and B. B. Friedman.

2 - Conveys a simple description completely and correctly.
Conveys the simple description completely and correctly, but elaborates and in so doing makes some error or errors of vocabulary, grammar, or pronunciation — errors which interfere little with the understandability of the utterance.

1 - Conveys the simple description with one or more errors of vocabulary, grammar or pronunciation, these errors being such as *not* to interfere with the understandability of the simple description.

0 - Conveys very little meaning.
Conveys the wrong meaning.
Makes errors which obscure the meaning.
Says nothing.

Part II, designated as Sustained Speech, calls for the student to discourse without interruption for two to three minutes on a topic assigned to him on the spot, such as the following:

Lower Level:

You are talking with a Spanish-speaking person who has never been to the United States. Describe to him the town or city in which you live. (If further stimulus is necessary before time is up:) This person is also interested in what a North American home looks like. Describe to him the home in which you (or your parents) live.

Upper Level:

You have met a young German in Europe who seems to you to have the makings of an outstanding American citizen. You resolve to try to convince him that he should emigrate to the United States.

Talk to him about the United States so that you may help him decide whether he would like to come. (If further stimulus is necessary:) Your young German friend is interested in American schools. Describe to him life at the school you attend.

In this part of the test maximum precision of rating seemed attainable through attention to four separate qualities of response; consequently the following rating scale was devised:

Fluency

2 - Speaks smoothly, phrasing naturally according to his thoughts.
1 - Occasionally hesitates in order to search for the right word or to correct an error.
0 - Speaks so haltingly that it is difficult to understand the thought he is conveying.

Vocabulary

2 - Vocabulary adequate for expressing the ideas he wishes to convey.
1 - Manages to convey his ideas in part, but in several instances uses an incorrect word or fails to find any word to use.
0 - Cannot communicate his thought because he does not have an adequate vocabulary.

Pronunciation and Enunciation

2 - Sufficiently approaches native speech to be completely understandable.
1 - Can be understood, though with difficulty, because there are sounds which he does not utter correctly.
0 - Would not be understood by natives because his pronunciation is so different from theirs.

Grammatical Correctness

2 - Speaks correctly with no serious errors in correct grammatical usage.
1 - Speech is understandable, but there are serious grammatical errors.
0 - Speech is not readily understandable because it is so full of grammatical errors.

Part III, the Conversation, consists of a directed exchange of remarks between the student and a native speaker whose voice is recorded on a phonograph disc. The student is asked to imagine that he is in the company of a friend whose native language is French, German or Spanish as the case may be. The friend speaks to him, and immediately afterward another voice on the record directs the student in English what to reply to his friend. For example, the friend may say: "¿Cómo está usted?", whereupon the English voice says: "Tell him you're fine and ask him how *he* is." Pauses are provided in the record while the student makes his contributions, which are rated by the examiner according to the following scale:

2 - Expresses idea accurately.

1 **- Partially incorrect.**
Conveys the correct idea but has one or more errors of grammar.
Conveys almost the correct idea, having one or two errors of vocabulary.

0 **- Only small part of idea conveyed.**
Wrong idea conveyed.
Not understandable.
No utterance made.

At the end of each pause, a buzzer warns the student to stop speaking and prepare for the next interchange.

Just as adding the uncontrollable factors inherent in an oral response seemed undesirable in the testing of aural comprehension, so in testing oral production we did not wish to assume that a stimulus spoken in the foreign language would necessarily be correctly understood. Hence, in the Picture Series the stimulus is purely visual. In the Sustained Speech it is merely the preliminary directions in English. In the Conversation the actual stimulus is the English voice telling the student what to say; the remarks of the foreign friend serve only to provide the illusion of a real conversation, but they do not have to be accurately understood before a correct response can be made.

TESTS OF WRITING

Though it was theoretically desirable that our testing program should be comprehensive, we did not include measures of writing for two reasons. One was that none of the courses collaborating most actively with us stressed this skill to any great degree. Even had this not been the case, the second reason would have been decisive, lack of testing time. To test reading, aural comprehension, and oral production required large blocks of class time, which was already far too scarce. Even the testing of oral ability, a much more important objective for all these programs, seemed likely to suffer from this lack, and events proved this fear to have been justified. To have attempted to measure writing skill also would merely have produced relatively unimportant data at the further expense of much more vital information.

Library

Modern Language Association
of America

Organized 1883

Chapter IV

The Aural and Oral Testing Program

A primary purpose of the Investigation was the appraisal of basic language courses of the currently prevalent "experimental" type. With few exceptions, these experimental programs had a common basis in both method and objective. The method was said to be "intensive," comprising six or more contact hours a week divided between demonstrations by the course director and drill sessions with a native-speaking guide, and using textual materials especially adapted to this kind of study. The aim was first to give the student skill in speaking and understanding the language and then to build a reading knowledge on the foundation of his oral and aural skills. We use the term "experimental" to denote a course embodying these essential features of aim and method, as contrasted with the "conventional" course. Thus the course which adds one or two hours of aural or oral practice to the three or four regular weekly classes, or the course which covers two or more terms' study in one by increasing the hours per week while not otherwise altering the method, does not constitute an "experimental" course as we define it.

AURAL COMPREHENSION

Since the aural comprehension tests described in Chapter III were newly developed, we had to have a standard which would render the scores meaningful. It seemed wisest to base this standard on the results achieved by conventional instruction and to measure those of experimental programs against them. We therefore solicited the co-operation of language teachers throughout the nation in establishing general norms for our tests. Teachers everywhere were invited, through publicity in the journals and at association meetings, to use the tests free of charge if they were interested in helping with this task and in measuring the aural attainments of their students. The response was considerable and the demand rose steadily between April, 1945, and July, 1947. During this period our aural tests were administered in 56 colleges and in 44 secondary schools, covering all sections of the United States.[1] In a few institutions located near headquarters, members of the Investigation staff were able themselves to administer the aural tests. The vast majority of administrations, however, were handled by the teachers themselves, with the phonograph records and supplementary materials (individual test-booklets, answer-sheets, special pencils) furnished to them free of charge.

Table I shows the total number of aural comprehension tests administered for the Investigation, distributed according to languages and levels and according to conventional and experimental groups.

Standardization of the Aural Tests

As soon as a sizeable body of test scores had accumulated the work of standardizing the aural tests was begun with the establishment of decile norms for part and total scores. Since the colleges are usually able to maintain relatively

---

[1] In addition, tests were given in three schools in Canada, and in one Mexican college to American exchange students.

TABLE I

Cumulative Totals of Aural Tests Administered

		LOWER LEVEL		UPPER LEVEL	
		NUMBER OF CASES	NUMBER OF INSTITUTIONS	NUMBER OF CASES	NUMBER OF INSTITUTIONS
	Conventional	2700	42	1724	24
French	Experimental	654	11	314	6
	Total	3354		2038	
	Conventional	1392	25	323	7
German	Experimental	531	6	332	5
	Total	1923		655	
	Conventional	90	8	20	3
Russian	Experimental	83	3	39	3
	Total	173		59	
	Conventional	5207	56	2804	38
Spanish	Experimental	686	12	478	7
	Total	5893		3282	

higher standards of language achievement than the high schools, two basic sets of norms were established: one for college students and another for students in secondary schools.

Since frequent mention will be made in the following pages of the norms for both the Co-operative Tests and our own, a reminder of the caution necessary in using them is probably not amiss. At their best the norms for any test indicate only the current level of achievement by certain classes of students. This degree of attainment may not be desirable or even satisfactory. Too frequently an instructor comparing the average of his class with national norms loses sight of this fact. If his group equals or surpasses the national mean or median, he is likely to conclude that he has been successful and that his students are learning all they need to know. (Yet no teacher would feel equally complacent merely because he learned that his salary equaled th average national income.) General average performance on an test may be scandalously low; and, whatever the level, teachers with superior students, better facilities, and other advantages should certainly set their sights higher than run-of-the-mill performance. Thus while norms serve as a yardstick against which to measure the performance of the experimental language classes, the scale should not be accepted uncritically. Even if the students of the newer programs prove to surpass average performance by a considerable amount, this fact is not in itself proof that these students have necessarily reached a satisfactory level of language proficiency.

(a) *College Norms*

As stated in Chapter III, the two levels of test in each language were designed to be used with students who had accumulated, respectively, totals of 130 or less hours and of 150 or more hours. However, as scores from the various colleges came in, it became evident that the limits set for the respective levels were neither sufficiently tight nor strictly enough observed by teachers using the tests. As a result, more than one norm-group had to be established for each level.

Norm-groups 1A, 1B, 2. At the lower level, Group 1 consisted of students tested while taking a beginning course and Group 2 was composed of those tested in an intermediate course. Group 1 was further subdivided as follows: the relatively small number of students who, when tested, had received less than 90 hours of instruction were designated as Group 1A, while the larger number of elementary students who had had 90 or more hours were called Group 1B.

Norm-groups 11, 12. At the upper level, Group 11 consisted of students tested while taking an intermediate course, and Group 12 was composed of those tested in an advanced (third-year or higher) course.

The norms for the lower and upper level forms of the French, German, and Spanish tests were based exclusively on scores made by students in conventional courses with classes meeting less than six hours per week. These norms could thus be utilized as an index of the aural ability of American language students generally, independent of that of students specifically taught this skill; and the performance of experimental students could be checked directly against the normal standard.

Table II sets forth the decile norms as established for the two levels in French, German, and Spanish. The Russian norms are presented separately at a later point.

TABLE II

Decile Norms for Aural Tests – Colleges*
French, German, Spanish

DEFINITION OF GROUPS (ALL LANGUAGES)

Lower Level
- 1A = Beginning course: less than 90 hours
- 1B = Beginning course: more than 90 hours
- 2 = Intermediate course

Upper Level
- 11 = Intermediate course
- 12 = Advanced course

---

*It will be recalled that the nine deciles are merely the lines which divide the continuum of distributed scores into ten equal portions. Thus, for example, the ninth decile divides the top tenth from the ninth tenth; the fifth decile divides the sixth tenth from the fifth tenth and thus the top half from the bottom half; the first decile divides the second tenth from the bottom tenth. The raw scores are given for the most part in class intervals of two, three or four, for the sake of compactness. Looking at the norm-table for French Lower Level, Form A, we find for example that in the total, the score-interval 24-27 is cut by the 7th decile for Group 1A, by the 5th decile for Group 1B, and by the 4th decile for Group 2. This means that 70 per cent of the students in Group 1A, 50 per cent of those in Group 1B, and 40 per cent of those in Group 1B made total scores of 24-27 or lower.

The fifth decile indicates the median or middle score, which in symmetrical distributions coincides with the mean or average score.

FRENCH — LOWER LEVEL

Form A Group 1A = 44 students in 3 colleges
Group 1B = 1808 students in 21 colleges
Group 2 = 196 students in 11 colleges

PART I = 25 ITEMS				PART II = 30 ITEMS				PART III = 41 ITEMS				TOTAL = 96 ITEMS			
RAW	DECILES			RAW	DECILES			RAW	DECILES			RAW	DECILES		
SCORE	1A	1B	2	SCORE	1A	1B	2	SCORE	1A	1B	2	SCORE	1A	1B	2
13			9	21-22			9	24-25	.		9	56-59			9
12				19-20		9	8	22-23				52-55			
11			8	17-18		8	7	20-21			8	48-51			
10		9	7	15-16			6	18-19		9		44-47			8
9	9			13-14	9	7	5	16-17	9	8	7	40-43		9	7
8	8			11-12	8	5	4	14-15	8		6	36-39		8	6
7		8	6	9-10	5	4	2	12-13		7	5	32-35	9		5
6	7	7	5	7-8	4	3		10-11	7	5	4	28-31	8	7	
5	6	6		5-6	3	2	1	8-9	5	4		24-27	7	5	4
4	5	5	4	3-4	2	1		6-7	4	3	3	20-23	5	4	2
3	4	4	3	1-2	1			4-5	3	2	2	16-19	4	3	
2	3	3	2					2-3	2	1	1	12-15	2	2	1
0	1	1	1					0-1	1			8-11	1	1	
Means	4.8	4.4	6.7	Means	8.4	11.6	14.3	Means	8.8	10.3	13.1	Means	21.9	26.4	33.9

Form B Group 1B = 71 students in 1 college
Group 2 = 54 students in 1 college

PART I = 25 ITEMS			PART II = 30 ITEMS			PART III = 35 ITEMS			TOTAL = 90 ITEMS		
RAW	DECILES		RAW	DECILES		RAW	DECILES		RAW	DECILES	
SCORE	1B	2	SCORE	1B	2	SCORE	1B	2	SCORE	1B	2
12	9	9	21-22	9	9	18-19		9	40-43	9	9
11			19-20			16-17	9	8	36-39	8	8
10	8	8	17-18	8	8	14-15		7	32-35		7
9			15-16	7	6	12-13			28-31	7	6
8	7	7	13-14	5	5	10-11	8	6	24-27	5	5
7	6	6	11-12	4	4	8-9	5	5	20-23	3	3
6	5	5	9-10	3	3	6-7	4	3	16-19	2	2
5	4	4	7-8	2	2	4-5	3	2	12-15	1	1
4	3	3	5-6	1	1	2-3	2	1			
3						0-1	1				
2	2	2									
1	1	1									
Means	6.6	6.9	Means	13.0	13.5	Means	8.0	10.5	Means	27.5	30.9

FRENCH — UPPER LEVEL

Form A Group 11 = 777 students in 14 colleges
Group 12 = 141 students in 7 colleges

PART I = 25 ITEMS			PART II = 40 ITEMS			PART III = 15 ITEMS			TOTAL = 80 ITEMS		
RAW	DECILES		RAW	DECILES		RAW	DECILES		RAW	DECILES	
SCORE	11	12	SCORE	11	12	SCORE	11	12	SCORE	11	12
22		9	36		9	11		9	65		9
20-21		8	33-35		8	10			60-64		
18-19		7	30-32		7	9		8	55-59		8
16-17	9	6	27-29	9		8	9	7	50-54		7
14-15	8	5	24-26	8	6	7		6	45-49	9	6
12-13	7	4	21-23	7	5	6	8	5	40-44	8	5
10-11	6	3	18-20	5	4	5			35-39	7	4
8-9	5	2	15-17	4	2	4	7	4	30-34	6	3
6-7	4		12-14	3	1	3	6	3	25-29	5	2
4-5	3	1	9-11	2		2	5	2	20-24	4	1
2-3	1		6-8	1		1	4		15-19	2	
						0	3	1	10-14	1	
Means	8.7	14.0	Means	17.8	23.6	Means	3.1	5.7	Means	29.4	43.0

Form B Group 11 = 344 students in 5 colleges
Group 12 = 104 students in 3 colleges

PART I = 25 ITEMS			PART II = 40 ITEMS			PART III = 15 ITEMS			TOTAL = 80 ITEMS		
RAW	DECILES		RAW	DECILES		RAW	DECILES		RAW	DECILES	
SCORE	11	12	SCORE	11	12	SCORE	11	12	SCORE	11	12
14-15		9	27-28		9	11		9	52-53		9
12-13		7	24-26		8	10		8	48-51		
10-11	9	6	21-23		7	9	9	7	44-47		8
8-9	8	5	18-20	9	6	8	8	6	40-43		
6-7	7	3	15-17	8	5	7		5	36-39		7
4-5	5	1	12-14	6	4	6	7	4	32-35	9	6
2-3	3		9-11	5	3	5	6		28-31	8	5
0-1	1		6-8	3	2	4	5	3	24-27	7	3
			3-5	2	1	3	4		20-23	6	2
						2	3	2	16-19	5	
						1			12-15	3	1
						0	2	1	8-11	1	
Means	5.2	9.6	Means	11.3	16.6	Means	4.3	6.6	Means	20.6	32.7

GERMAN – LOWER LEVEL

Group 1A = 300 students in 3 colleges
Group 1B = 473 students in 13 colleges
Group 2 = 171 students in 6 colleges

PART I = 25 ITEMS				PART II = 25 ITEMS				PART III = 36 ITEMS				TOTAL = 86 ITEMS			
RAW	DECILES			RAW	DECILES			RAW	DECILES			RAW	DECILES		
SCORE	1A	1B	2	SCORE	1A	1B	2	SCORE	1A	1B	2	SCORE	1A	1B	2
14-15		9	9	16		9	9	22-23		9		48-49			9
12-13	9	8	8	14-15				20-21			9	44-47		9	
10-11			7	12-13	9	8	8	18-19			8	40-43			
8-9	8	7	5	10-11		7	7	16-17	9	8		36-39		8	8
6-7	7	5	4	8-9	8	6	6	14-15			7	32-35	9		7
4-5	5	4	3	6-7	7	5	5	12-13		7	6	28-31		7	6
2-3	4	3	2	4-5	5	3	3	10-11	8	6		24-27			
0-1	3	1	1	2-3	4	2	2	8-9	7	5	5	20-23	8	6	5
				0-1	2	1	1	6-7	6	4	4	16-19	7	5	4
								4-5	5	3	3	12-15	5	3	3
								2-3	3	2	2	8-11	4	2	2
								0-1	2	1	1	4-7	2	1	1
												0-3	1		
Means	4.7	6.8	8.3	Means	4.7	7.3	8.1	Means	6.7	9.6	10.6	Means	15.8	23.4	26.5

GERMAN – UPPER LEVEL

Group 11-12 = 84 students in 3 colleges

PART I = 25 ITEMS		PART II = 40 ITEMS		PART III = 15 ITEMS		TOTAL = 80 ITEMS	
RAW SCORE	DECILES	RAW SCORE	DECILES	RAW SCORE	DECILES	RAW SCORE	DECILES
12-13	9	30	9	10	9	52-53	9
10-11	8	27-29		9	8	48-51	
8-9	7	24-26		8	7	44-47	
6-7	5	21-23	8	7	6	40-43	8
4-5	4	18-20	7	6	5	36-39	
2-3	3	15-17	6	5		32-35	
0-1	1	12-14	5	4	4	28-31	7
		9-11	3	3	2	24-27	6
		6-8	2	2	1	20-23	5
		3-5	1			16-19	3
						12-15	2
Mean	5.9	Mean	15.2	Mean	6.0	Mean	26.9

SPANISH — LOWER LEVEL

Form A Group 1A = 31 students in 2 colleges
Group 1B = 1412 students in 22 colleges
Group 2 = 563 students in 15 colleges

PART I = 25 ITEMS

RAW	DECILES		
SCORE	1A	1B	2
16			9
15			
14			
13			
12		9	8
11			
10			7
9		8	
8	9		6
7	8	7	
6	7	6	5
5	6	5	4
4	5	4	3
3	4		
2	3	3	2
1	2	2	1
0	1	1	
Means	4.4	5.6	7.6

PART II = 25 ITEMS

RAW	DECILES		
SCORE	1A	1B	2
20			9
19			
18			8
17		9	
16			7
15			
14		8	6
13	9	7	
12		6	5
11			4
10	8	5	3
9		4	
8	7	3	2
7	6		
6	5	2	1
5	4		
4	3	1	
Means	7.2	10.5	12.9

PART III = 38 ITEMS

RAW	DECILES		
SCORE	1A	1B	2
26-27			9
24-25			
22-23		9	8
20-21			7
18-19		8	
16-17	9	7	6
14-15	7	6	5
12-13	6	5	4
10-11	5	4	3
8-9	3	3	2
6-7	2	2	
4-5	1	1	1
Means	11.5	13.4	15.7

TOTAL = 88 ITEMS

RAW	DECILES		
SCORE	1A	1B	2
60-63			9
56-59			
52-55			
48-51			8
44-47		9	
40-43			7
36-39		8	6
32-35	9	7	5
28-31		5	4
24-27	8	4	3
20-23	5	3	2
16-19	3	2	1
12-15	1	1	
Means	22.9	29.4	35.8

Form B Group 1B = 141 students in 4 colleges
Group 2 = 165 students in 3 colleges

PART I = 25 ITEMS

RAW	DECILES	
SCORE	1B	2
18		9
16-17		
14-15		8
12-13	9	7
10-11	8	5
8-9	7	4
6-7	5	3
4-5	4	2
2-3	2	1
0-1	1	
Means	6.7	10.1

PART II = 25 ITEMS

RAW	DECILES	
SCORE	1B	2
18-19		9
16-17	9	7
14-15	8	6
12-13	7	5
10-11	6	4
8-9	5	3
6-7	3	2
4-5	2	1
2-3	1	
Means	9.4	12.1

PART III = 38 ITEMS

RAW	DECILES	
SCORE	1B	2
28		9
26-27		
24-25		8
22-23	9	7
20-21	8	
18-19	7	6
16-17	6	5
14-15	5	4
12-13	4	
10-11	3	3
8-9	2	2
6-7	1	
4-5		1
Means	14.4	16.5

TOTAL = 88 ITEMS

RAW	DECILES	
SCORE	1B	2
60-63		9
56-59		
52-55		8
48-51		7
44-47	9	
40-43	8	6
36-39	7	5
32-35	6	4
28-31	5	3
24-27	4	
20-23	3	2
16-19	2	1
12-15	1	
Means	30.3	38.7

SPANISH — UPPER LEVEL

Form A Group 11 = 1044 students in 19 colleges
Group 12 = 302 students in 10 colleges

PART I = 25 ITEMS

RAW SCORE	DECILES 11	DECILES 12
20		9
18-19		
16-17	9	8
14-15		7
12-13	8	5
10-11	6	4
8-9	5	3
6-7	4	2
4-5	2	1
2-3	1	
Means	9.4	11.5

PART II = 45 ITEMS

RAW SCORE	DECILES 11	DECILES 12
34-35		9
32-33	9	
30-31		8
28-29	8	7
26-27	7	6
24-25	6	5
22-23		
20-21	5	4
18-19	4	3
16-17	3	2
14-15	2	1
12-13		
10-11	1	
Means	21.7	24.3

PART III = 15 ITEMS

RAW SCORE	DECILES 11	DECILES 12
9	9	9
8	8	
7		8
6	7	7
5	6	6
4	5	5
3	4	4
2	3	3
1	2	2
0	1	1
Means	4.3	4.4

TOTAL = 85 ITEMS

RAW SCORE	DECILES 11	DECILES 12
60-64		9
56-59		
52-55	9	8
48-51		
44-47	8	7
40-43	7	6
36-39	6	5
32-35	5	4
28-31	4	3
24-27	2	2
20-23	1	1
Means	35.2	40.1

Form B Group 11 = 303 students in 3 colleges
Group 12 = 127 students in 4 colleges

PART I = 25 ITEMS

RAW SCORE	DECILES 11	DECILES 12
18-19		9
16-17		7
14-15	9	5
12-13	8	4
10-11	5	2
8-9	4	
6-7	3	1
4-5	2	
2-3	1	
Means	9.3	13.5

PART II = 44 ITEMS

RAW SCORE	DECILES 11	DECILES 12
26		9
24-25		
22-23		8
20-21		
18-19		7
16-17	9	6
14-15		5
12-13	8	4
10-11	7	3
8-9	6	2
6-7	5	
4-5	4	1
2-3	2	
0-1	1	
Means	8.1	14.7

PART III = 15 ITEMS

RAW SCORE	DECILES 11	DECILES 12
14		9
13		
12		8
11		
10	9	7
9	8	5
8	7	4
7	5	
6		3
5	4	
4	3	2
3		
2	2	1
1		
0	1	
Means	6.0	8.2

TOTAL = 84 ITEMS

RAW SCORE	DECILES 11	DECILES 12
56		9
52-55		
48-51		
44-47		8
40-43		6
36-39	9	5
32-35	8	4
28-31	7	3
24-27	6	2
20-23	5	
16-19	3	1
12-15	2	
8-11	1	
Means	23.9	36.2

Table III shows the mean scores for both parts and totals made by all experimental students tested, as contrasted with the respective mean scores of the two norm-groups with which they are most nearly comparable. Such comparison excludes Group 1A at the lower level, since all experimental students had accumulated at least 90 hours when tested. Scores of the experimental groups are listed between those of the two norm-groups, since comparisons should be made in both directions: i.e., at the lower level, the experimental groups are almost exclusively composed of students taking a beginning course and are in this respect comparable to Group 1B; at the same time many of them have received as many instructional hours as students in intermediate courses and are thus comparable also to Group 2. The same considerations apply, *mutatis mutandis*, at the upper level.

In most cases the means of the experimental groups fall between those of the two norm-groups. Exceptions favor Group 1B over the experimental in Lower-Level French (both forms); the experimental over Group 12 in Upper-Level French (Form B only) and in Upper-Level Spanish (Form A only). In any case, these averages for all the experimental groups combined are provided only as a very rough indication of the achievement of experimentation in general. They are derived from groups which are widely divergent in total hours accumulated, in objectives, teaching procedure and the rest; and hence no attempt has been made to determine whether any of the observed differences are significant — a procedure reserved for the study of more homogeneous groups in particular programs.

Russian. The aural tests in Russian were not used widely enough to warrant the establishment of more than a single set of norms at each level, both conventional and experimental groups. Table IV shows the distribution for each level. Only the means, reported at the bottom of the columns, distinguish between conventional and experimental students.

TABLE III

Mean Scores of Norm-Groups and Experimental Groups

	PART I MEANS	PART II MEANS	PART III MEANS	TOTAL MEANS
French — Lower Level				
Form A				
Norm-group 1B (N=1808)	4.4	11.6	10.3	26.4
Experimental groups (N=498)	5.0	11.8	9.4	26.2
Norm-group 2 (N=196)	6.7	14.3	13.1	33.9
Form B				
Norm-group 1B (N=71)	6.6	13.0	8.0	27.5
Experimental groups (N=141)	6.2	12.3	7.1	25.7
Norm-group 2 (N=54)	6.9	13.5	10.5	30.9
French — Upper Level				
Form A				
Norm-group 11 (N=777)	8.7	17.8	3.1	29.4
Experimental groups (N=151)	12.1	20.4	4.3	36.8
Norm-group 12 (N=141)	14.0	23.6	5.7	43.0
Form B				
Norm-group 11 (N=344)	5.2	11.3	4.3	20.6
Experimental groups (N=163)	8.3	17.4	7.3	33.0
Norm-group 12 (N=104)	9.6	16.6	6.6	32.7
German — Lower Level				
Norm-group 1B (N=473)	6.8	7.3	9.6	23.4
Experimental groups (N=480)	7.8	7.2	8.5	23.6
Norm-group 2 (N=171)	8.3	8.1	10.6	26.5
German — Upper Level				
Norm-group 11-12 (N=84)	5.9	15.2	6.0	26.9
Experimental groups (N=317)	6.6	18.6	6.4	31.5
Spanish — Lower Level				
Form A				
Norm-group 1B (N=1412)	5.6	10.5	13.4	29.4
Experimental groups (N=550)	5.6	10.6	13.4	29.5
Norm-group 2 (N=563)	7.6	12.9	15.7	35.8
Form B				
Norm-group 1B (N=141)	6.7	9.4	14.4	30.3
Experimental groups (N=136)	10.5	11.7	12.5	34.8
Norm-group 2 (N=165)	10.1	12.1	16.5	38.7
Spanish — Upper Level				
Form A				
Norm-group 11 (N=1044)	9.4	21.7	4.3	35.2
Experimental groups (N=291)	13.0	26.0	5.0	44.0
Norm-group 12 (N=302)	11.5	24.3	4.4	40.1
Form B				
Norm-group 11 (N=303)	9.3	8.1	6.0	23.9
Experimental groups (N=187)	12.0	10.3	6.4	28.7
Norm-group 12 (N=127)	13.5	14.7	8.2	36.2

TABLE IV

Decile Norms for Russian Aural Tests

LOWER LEVEL

157 students in 10 colleges

PART I = 25 ITEMS		PART II = 25 ITEMS		PART III = 36 ITEMS		TOTAL = 86 ITEMS	
RAW SCORE	DECILES	RAW SCORE	DECILES	RAW SCORE	DECILES	RAW SCORE	DECILES
20-21	9	22	9	30-32	9	72	9
18-19	8	20-21		27-29	8	66-71	
16-17		18-19	8	24-26		60-65	8
14-15		16-17	7	21-23	7	54-59	
12-13	7	14-15	6	18-20	6	48-53	7
10-11	6	12-13	5	15-17	5	42-47	6
8-9	5	10-11	4	12-14	3	36-41	5
6-7	4	8-9	2	9-11	2	30-35	4
4-5	3	6-7	1	6-8		24-29	3
2-3	2			3-5	1	18-23	2
0-1	1					12-17	1
Means	9.8		13.3		17.9		40.8
Means, Conventional (N=90)*	11.4		14.7		19.1		45.2
Means, Experimental (N=83)	8.2		11.9		15.8		35.9

UPPER LEVEL

59 students in 6 colleges

PART I = 25 ITEMS		PART II = 34 ITEMS		PART III = 15 ITEMS		TOTAL = 74 ITEMS	
RAW SCORE	DECILES	RAW SCORE	DECILES	RAW SCORE	DECILES	RAW SCORE	DECILES
18	9	30-31	9	14-15	9	60-62	9
16-17	8	28-29	8	12-13	8	55-59	8
14-15	7	26-27	7	10-11	7	50-54	7
12-13	5	24-25	6	8-9	5	45-49	6
10-11	4	22-23	5	6-7	4	40-44	5
8-9	3	20-21	4	4-5	3	35-39	4
6-7	2	18-19		2-3	2	30-34	3
4-5	1	16-17	3	0-1	1	25-29	2
		14-15	2			20-24	
		12-13				15-19	1
		10-11	1				
Means	11.2		21.7		8.0		41.1
Means, Conventional (N=20)	13.7		24.0		9.9		47.6
Means, Experimental (N=39)	9.9		20.4		6.9		37.2

*Includes 16 additional students who had accumulated more than 130 hours.

(b) *Secondary School Norms*

Secondary-school students on each level were classified into norm-groups according to the number of semesters completed (or nearly so) at the time of testing. Decile norms were established for French and Spanish at both levels (Form A only), and in German at the lower level, as set forth in Table V.

TABLE V

*Decile Norms for Aural Tests – Secondary Schools**
French, German, Spanish

Definition of Groups (all languages)

Lower Level
OA = one semester
OB = two semesters (one year)
OC = three semesters
OD = four semesters (two years)
OE = five semesters
OF = six semesters (three years)

Upper Level
10D = four semesters (two years)
10E = five semesters
10F = six semesters (three years)
10G = seven semesters
10H = eight semesters (four years)

---

*For interpretation of decile tables see note, p. 65.

FRENCH — LOWER LEVEL

Form A Group OA = 28 students in 2 schools
Group OB = 305 students in 8 schools
Group OC-D = 98 students in 5 schools

PART I = 25 ITEMS				PART II = 30 ITEMS				PART III = 41 ITEMS				TOTAL = 96 ITEMS			
RAW	DECILES			RAW	DECILES			RAW	DECILES			RAW	DECILES		
SCORE	OA	OB	OC-D	SCORE	OA	OB	OC-D	SCORE	OA	OB	OC-D	SCORE	OA	OB	OC-D
9	9			18-19		9	9	18-19			9	36-39			9
8	8		9	16-17			8	16-17	9			32-35	9	9	8
7		9	8	14-15		8	7	14-15		9	8	28-31			7
6		8	7	12-13	9		6	12-13	8		7	24-27	8	8	5
5		7	6	10-11	7	7	5	10-11	7	8	6	20-23	7	7	4
4	7	5	5	8-9	6	5	3	8-9	6	7	5	16-19	5	5	3
3	6			6-7	5	4	2	6-7		5	4	12-15	4	4	2
2	5	4	4	4-5	3	3	1	4-5	5	4	2	8-11	3	2	1
1	4	3	2	2-3	2	1		2-3	3	3		4-7	1	1	
0	3	2	1	0-1	1			0-1	1	1	1				
Means	3.4	3.5	4.0	Means	7.4	9.2	11.5	Means	7.2	6.8	9.8	Means	17.9	19.1	25.1

FRENCH - UPPER LEVEL

Form A Group 10D = 72 students in 3 schools
Group 10F = 30 students in 3 schools

PART I = 25 ITEMS			PART II = 40 ITEMS			PART III = 15 ITEMS			TOTAL = 80 ITEMS		
RAW	DECILES		RAW	DECILES		RAW	DECILES		RAW	DECILES	
SCORE	10D	10F	SCORE	10D	10F	SCORE	10D	10F	SCORE	10D	10F
16		9	24-25		9	6	9	9	42-44		9
14-15			22-23	9	8	5		8	39-41		
12-13		8	20-21	8	7	4	8		36-38		8
10-11	9	7	18-19	7	6	3	6	7	33-35	9	7
8-9	8	6	16-17	6	5	2	4	5	30-32	8	6
6-7	7	5	14-15	5	4	1	3	3	27-29	7	5
4-5	5	3	12-13	4	3	0	2	2	24-26	6	4
2-3	4	1	10-11	3	2				21-23	5	3
0-1	2		8-9						18-20	4	2
			6-7	2	1				15-17	3	
									12-14		1
									9-11	2	
Means	4.9	8.5	Means	14.4	17.3	Means	2.9	2.9	Means	21.9	28.6

GERMAN — LOWER LEVEL

Group OB = 160 students in 6 schools
Group OC = 32 students in 1 school
Group OD = 32 students in 4 schools

PART I = 25 ITEMS				PART II = 25 ITEMS				PART III = 36 ITEMS				TOTAL = 86 ITEMS			
RAW	DECILES			RAW	DECILES			RAW	DECILES			RAW	DECILES		
SCORE	OB	OC	OD	SCORE	OB	OC	OD	SCORE	OB	OC	OD	SCORE	OB	OC	OD
16			9	14			9	26			9	45-47			9
15				13	9			24-25				42-44	9		
14	9	9	8	12		9	8	22-23				39-41			8
13			7	11			7	20-21				36-38			
12			6	10			6	18-19	9			33-35			7
11				9				16-17				30-32			6
10			5	8	8			14-15			8	27-29	8	9	
9	8			7		8	5	12-13	8			24-26			5
8		8		6	7		4	10-11		9	6	21-23			
7	7		4	5	6	7	3	8-9	7	8	5	18-20	7	8	4
6		7	3	4	5	6	2	6-7	6	7	4	15-17	6		
5	6			3				4-5	5	6		12-14	5	7	3
4	5		2	2	4	5		2-3	4	5	3	9-11	4	6	2
3				1	3	4	1	0-1	2	1	2	6-8	3	5	1
2	4	6	1	0	2	2						3-5	1	4	
1	3	5										0-2		1	
0	2	4													
Means	5.3	3.9	9.4	Means	5.0	4.1	8.1	Means	7.4	5.3	10.0	Means	17.5	13.1	27.3

SPANISH — LOWER LEVEL

Form A

Group OA = 355 students in 5 schools
Group OB = 743 students in 11 schools
Group OC = 158 students in 3 schools
Group OD = 365 students in 10 schools
Group OE = 27 students in 2 schools
Group OF = 36 students in 3 schools

PART I = 25 ITEMS

RAW SCORE	DECILES OA	OB	OC	OD	OE	OF
20						9
19						8
18						
17					9	
16						7
15						6
14				9		5
13						4
12					8	3
11						2
10			9	8	7	1
9						
8	9	9	8	7		
7		8		6	6	
6	8	7	7	5	5	
5	7		6			
4	6	6	5	4	4	
3	5	5	4	3	3	
2	4	4	3	2	2	
1	3	3	2			
0	2	2	1	1	1	
Means	3.6	3.8	4.9	6.4	7.3	14.8

PART II = 25 ITEMS

RAW SCORE	DECILES OA	OB	OC	OD	OE	OF
22						9
21						
20						8
19						7
18					9	6
17						
16				9	8	5
15						4
14			9	8	7	3
13					6	2
12	9	9	8	7		
11				6		1
10	8	8	7	5	5	
9			6		4	
8	7	7	5	4	3	
7	6	6	4	3	2	
6	5	5	3	2		
5	4	4	2		1	
4	3	3		1		
3			1			
2	2	2				
1	1	1				
Means	6.5	6.5	8.4	10.4	11.7	17.1

SPANISH — LOWER LEVEL

(Form A)

(continued)

PART III = 38 ITEMS

RAW			DECILES			
SCORE	OA	OB	OC	OD	OE	OF
28-29						9
26-27						8
24-25					9	7
22-23				9		5
20-21			9			
18-19			8	8	8	
16-17	9	9			7	4
14-15		8	7	7	6	2
12-13	8		6	6	5	
10-11	7	7	5	5	4	
8-9	5	6	3	3	3	1
6-7	4	5		2		
4-5	3	4	2		2	
2-3	2	3	1	1	1	
0-1	1	2				
Means	8.8	8.1	11.9	12.0	13.7	20.2

TOTAL = 88 ITEMS

RAW			DECILES			
SCORE	OA	OB	OC	OD	OE	OF
68-71						9
64-67						8
60-63						
56-59						7
52-55						5
48-51					9	
44-47				9		4
40-43				8	8	3
36-39			9		7	2
32-35			8	7	6	1
28-31	9	9	7	6	5	
24-27	8	8	5	5	4	
20-23	7	7	4	3	3	
16-19	5	5	3	2	2	
12-15	4	4	2	1	1	
8-11	2	2	1			
4-7	1	1				
Means	18.7	18.2	24.9	28.3	32.3	51.9

SPANISH -- UPPER LEVEL

(Form A)

Group 10C = 61 students in 3 schools
Group 10D = 418 students in 6 schools
Group 10E = 63 students in 2 schools
Group 10F = 140 students in 8 schools
Group 10G-H = 47 students in 3 schools

PART I = 25 ITEMS

RAW SCORE	DECILES 10C	10D	10E	10F	10GH
18					9
17					
16				9	8
15					
14					
13		9		8	7
12	9			7	
11					
10	8	8	9	6	6
9	7		8		5
8		7	7	5	4
7	6	5	6	3	
6	5	4	5	2	
5	4		4		3
4	3	3	3	1	
3		2			2
2	2	1	2		1
1	1				
0			1		
Means	7.1	7.1	6.3	9.5	9.8

PART II = 45 ITEMS

RAW SCORE	DECILES 10C	10D	10E	10F	10GH
34-35					9
32-33					8
30-31					
28-29				9	7
26-27					6
24-25	9	9	9	8	5
22-23	8		8		
20-21	7	8	7	7	4
18-19	6	7	6	6	3
16-17	5	6	5	5	
14-15	4	5	4	4	2
12-13	3	4	3	3	1
10-11			2	2	
8-9	2	3	1	1	
6-7		2			
4-5	1				
2-3		1			
Means	16.1	14.0	17.2	17.5	24.5

PART III = 15 ITEMS

RAW SCORE	DECILES 10C	10D	10E	10F	10GH
10					9
9					
8		9		9	8
7					
6	9	8	9	8	7
5				7	6
4	8	7	8	5	5
3	7	5	5	4	4
2	5	4	4	2	3
1		3	3		2
0	4	2	2	1	1
Means	2.3	3.1	3.1	4.1	4.7

TOTAL = 85 ITEMS

RAW SCORE	DECILES 10C	10D	10E	10F	10GH
56-58					9
52-55					8
48-51					
44-47				9	7
40-43				8	6
36-39	9	9	9	7	5
32-35	8	8		6	4
28-31	7	7	8	5	3
24-27	5	6	5	4	2
20-23	4	5	3	2	1
16-19	3	3	1	1	
12-15	2	2			
8-11	1	1			
Means	25.4	24.1	26.5	30.9	37.8

Next to no aural testing was done in Upper-Level German or in Russian in secondary schools. Table VI lists the mean scores obtained on the Upper-Level German and on the Lower-Level Russian tests:

TABLE VI

Mean Scores for German (Upper Level) and Russian (Lower Level) in Secondary Schools

	PART I	PART II	PART III	TOTAL
German — Upper Level				
Mean, two schools (N=23)	6.9	14.2	5.1	25.8
Russian — Lower Level				
Mean, one school (N=11)	3.4	9.6	11.3	24.3

EVALUATION OF GENERAL AURAL ACHIEVEMENT

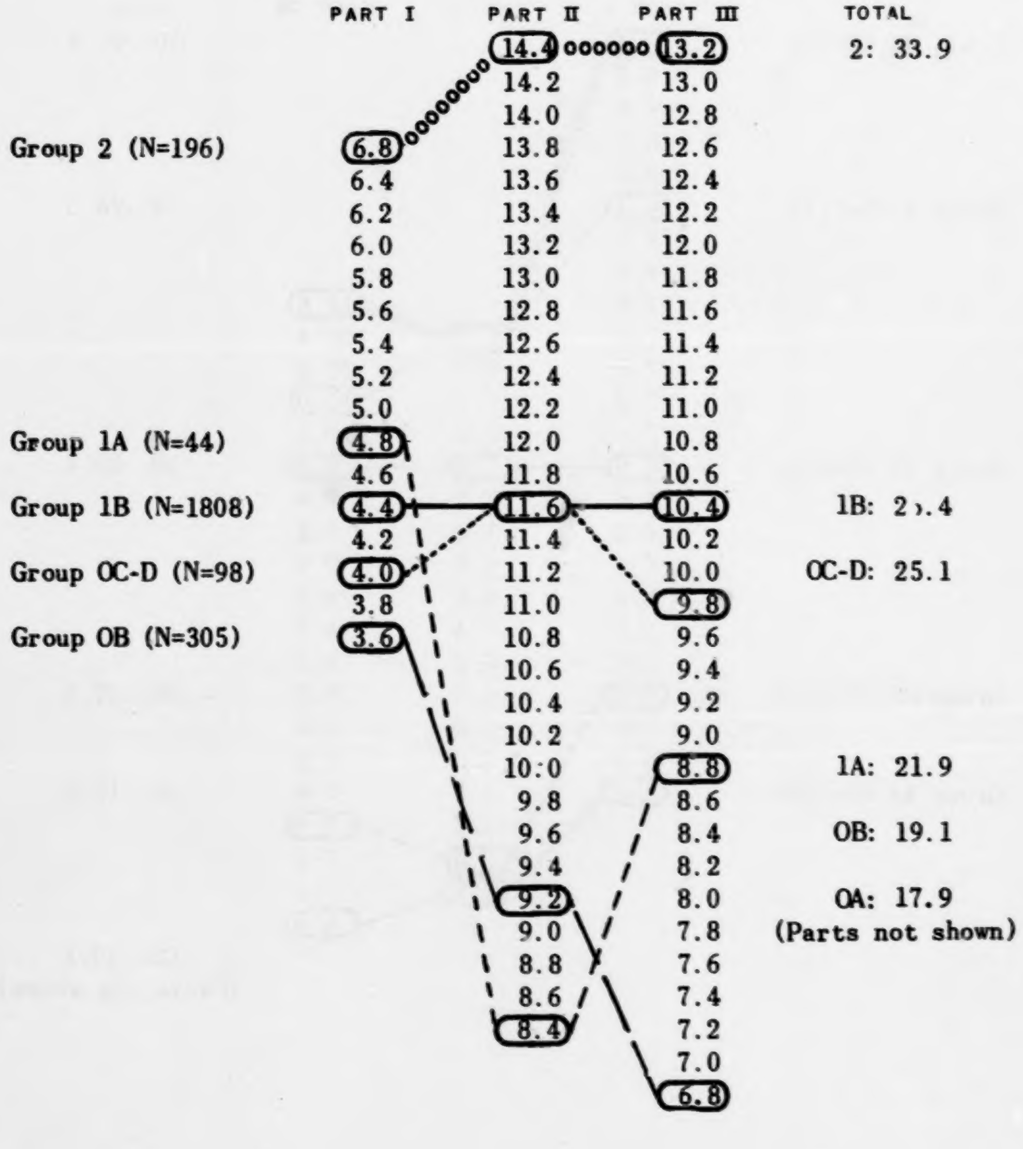
Achievement in Relative Terms

Before we attempt to discuss and interpret the results of the over-all aural testing program in any absolute terms, it may be helpful to plot the mean scores of the principal norm-groups, in order to point up certain interrelations not only of total but of part scores. For each level of French, German and Spanish, separate diagrams follow in which the part scores of the predominant norm-group (Group 1B for the lower level; Group 11 for the upper level) are plotted on a straight ("normalized") base-line. Respective part scores for the other norm-groups are plotted according to their deviation (in raw-score points) from this base-line. Total scores for all groups appear in descending order at the right of each diagram, those not plotted being included in parentheses.

French Lower Level: *(See chart on page 81.)*

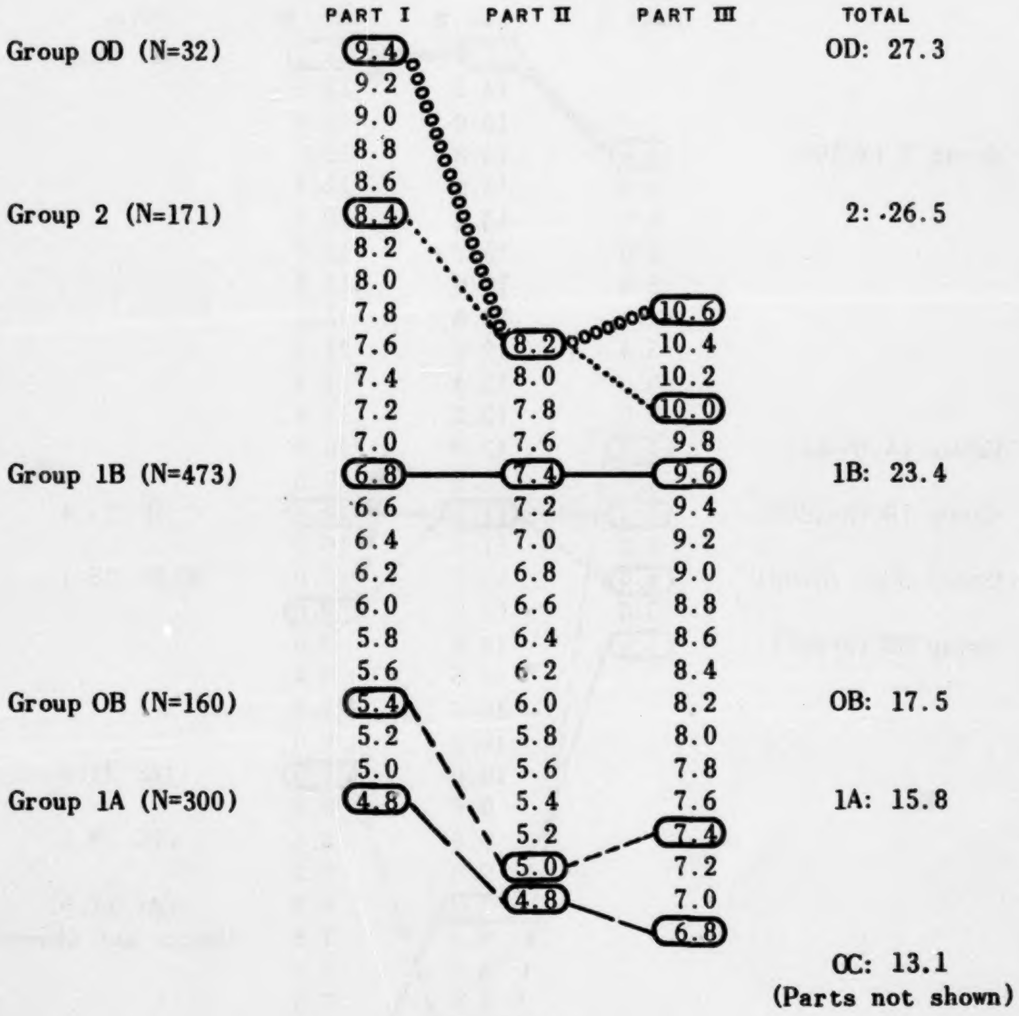
1. The tendency for all groups to be closest together on Part I is observable in nearly all the tests, and may be due

FRENCH - LOWER LEVEL (Form A)

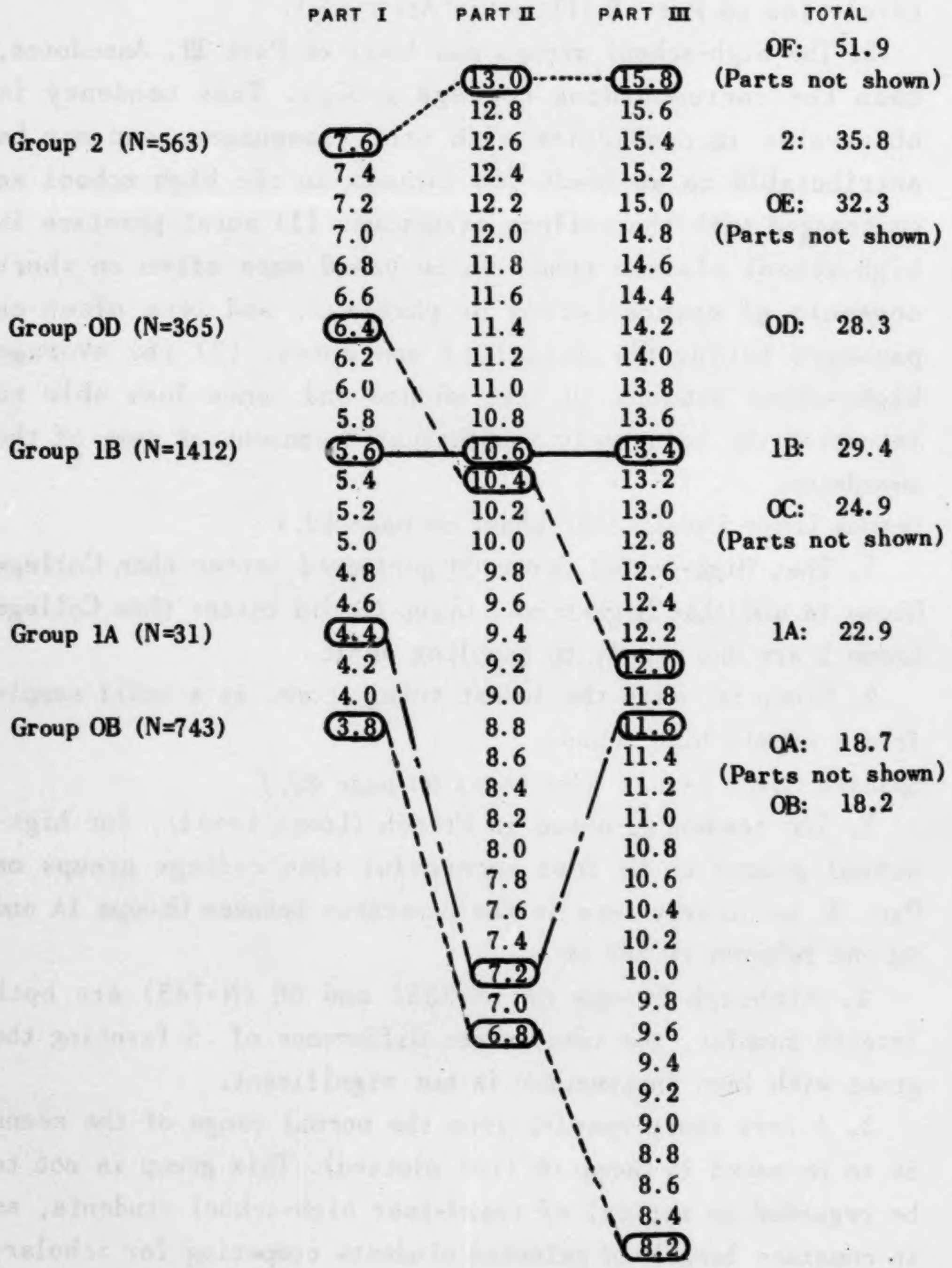


to psychological factors. Nervousness and confusion at the beginning of the test may well have lowered the scores of many students capable of better response, and may thus have reduced the spread at this point. (That Group 1A is actually higher than 1B on Part I is probably attributable to the fact that 1A represents a very small sample as compared with 1B.)

GERMAN - LOWER LEVEL



SPANISH — LOWER LEVEL (Form A)



2. Students tested before the end of the first year, whether in high school (OB) or in college (1A) are comparatively low on Part II (Phonetic Accuracy).

3. The high-school groups run lower on Part III, Anecdotes, than the corresponding college groups. This tendency is observable in connection with other languages, and may be attributable to at least two factors in the high school as contrasted with the college situation: (1) aural practice in high-school classes tends to be based more often on short segments of speech (words or phrases), and less often on passages having the length of anecdotes; (2) the average high-school student is less mature and hence less able to interpret the relatively sophisticated content of some of the anecdotes.

German Lower Level: *(See chart on page 82.)*

1. That High-school Group OB performed better than College Group 1A and that High-school Group OD did better than College Group 2 are due merely to sampling error.

2. Group OC, with the lowest total score, is a small sample from a single high school.

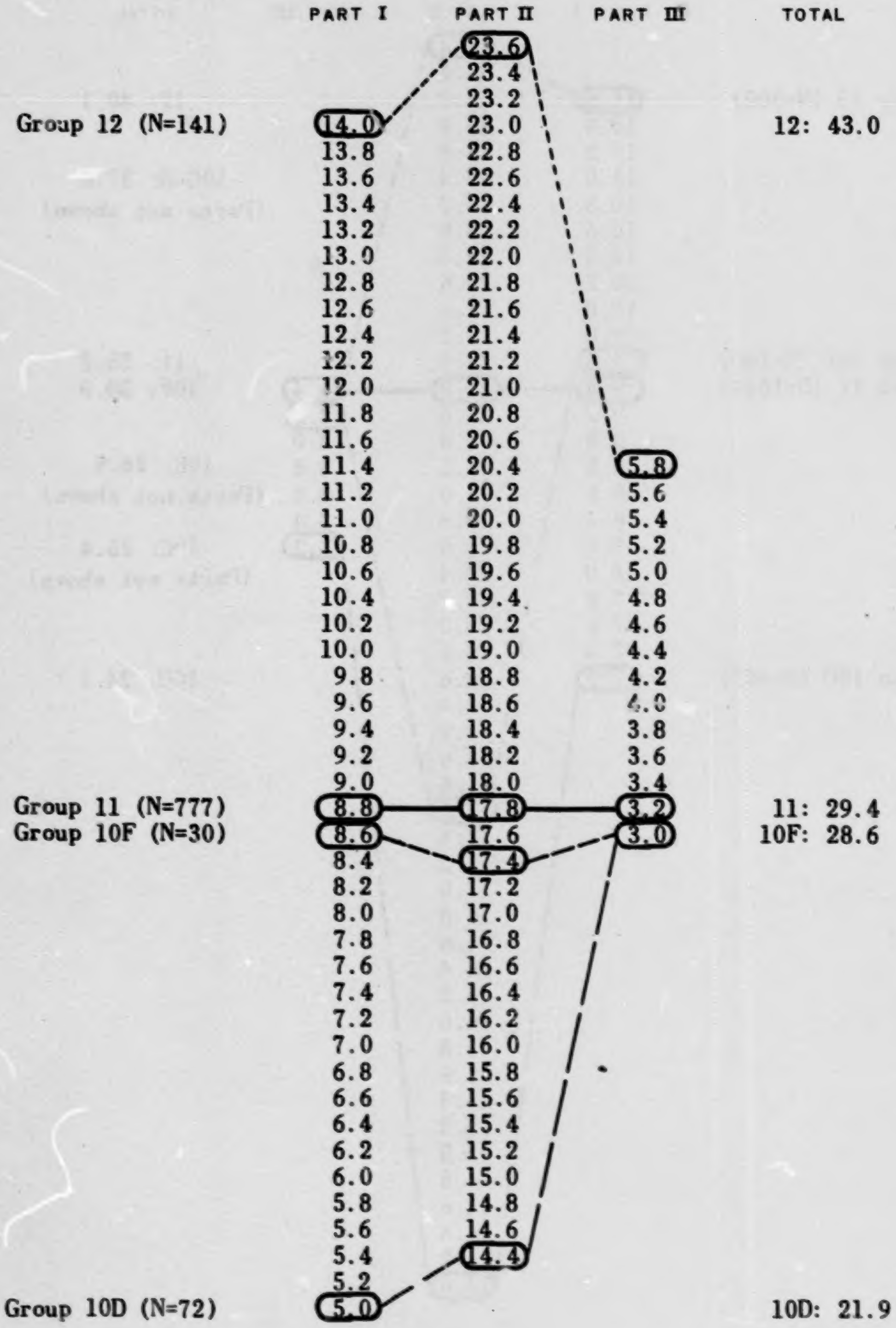
Spanish Lower Level: *(See chart on page 83.)*

1. The tendency, noted in French (Lower Level), for high-school groups to be less successful than college groups on Part III is clearly seen in the contrasts between Groups 1A and OB and between 1B and OD.

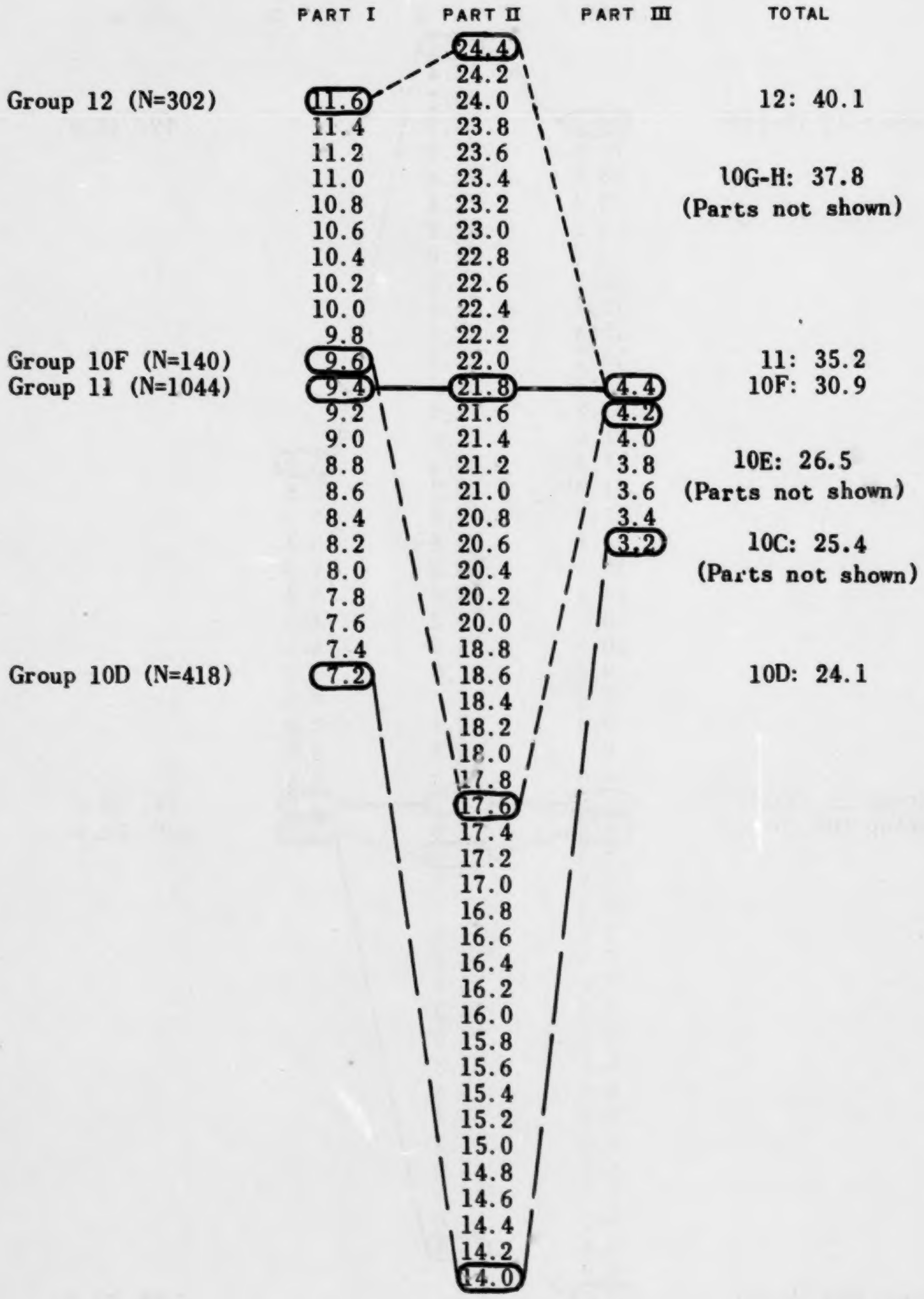
2. Although Groups OA (N=355) and OB (N=743) are both largish samples, the total-score difference of .5 favoring the group with less instruction is not significant.

3. A very sharp upswing from the normal range of the means is to be noted in Group OF (not plotted). This group is not to be regarded as typical of third-year high-school students, as it consists largely of selected students competing for scholarships.

FRENCH - UPPER LEVEL (Form A)



SPANISH - UPPER LEVEL (Form A)



French Upper Level: *(See chart on page 85.)*

1. High-school Group 10F (three years) and College Group 11 (intermediate), which closely parallel each other, show their maximum divergence in Part II (Anecdotes). This same tendency has already been noted in regard to the lower level.

2. Part III (Dialogue) is not only the most difficult section of the test to handle, but also has a small range of only 15 points. With the exception of the advanced college group, this part failed to spread the groups appreciably.

Spanish Upper Level: *(See chart on page 86.)*

1. The tendency for the high-school groups, as contrasted with the college groups, to make low scores on Part II (Anecdotes) is seen the most clearly in this picture.

2. As in French (Upper Level), Part III failed to produce any noteworthy spread.

We may now sum up our observations of main tendencies as follows:

1. At the lower level, High-school Group OD (two years) and College Group 1B (one year) approximate the same level of achievement. One year of high school seems not to produce as high results, however, as one semester of college (French, Spanish).

2. At the upper level, High-school Group 10F (three years) and College Group 11 (two years) are close to the same level.

3. At both levels, most groups show comparatively little spread in Part I, perhaps because of failure on the part of the more nervous students to respond at full capacity.

4. At both levels, high-school groups are less successful on the anecdote sections than college groups, perhaps because greater maturity is required to understand some of the anecdotes and because they have had less experience with long passages.

Achievement in Absolute Terms

In order to see the over-all aural achievement of the American language student, as measured by the Investigation tests, we may examine the results for important norm-groups, in terms of percentage of correct response. At the lower level we will most profitably consider the attainments of two groups, students who have completed one year, and two years of college study.[2] Within the corresponding norm-groups (1B and 2), the eighth decile indicates the score at or below which eighty per cent of the students stood, while the fifth decile will indicate the median score, which fifty per cent of the students failed to exceed. In Tables VII-1 and VII-2 we set opposite each score the percentage of the items answered correctly by the students making these scores.

Achievement at both the eighth and the fifth decile was generally higher on the upper-level than on the lower-level tests. Although it is possible that the lower-level tests were relatively harder tests, there is nevertheless good reason to see in these figures a proof that aural skill is acquired slowly and increases toward the end of two years' study. This conclusion is further confirmed by the notably higher percentages attained by the advanced (third-year) groups in French and Spanish.

Certain comparisons which we are about to make among performances in different languages are subject to extremely cautious interpretation. While identical levels of difficulty in all languages were aimed at in the construction of the tests at each level, there is no ground for assuming that results can be considered any more than approximately comparable. Further variables in this situation are the unmeasured differences, in background, intelligence, and the like.

---

[2] High-school groups are not included in the present data; the interested reader can make his own comparisons for any groups by consulting the high-school norms or the charts on pages 81-86.

TABLE VII-1

Percentages attained at 8th and 5th deciles

Lower Level

French (Form A)		Part I Definitions MAXIMUM POSSIBLE 25		Part II Phonetic Accuracy MAXIMUM POSSIBLE 30		Part III Anecdotes MAXIMUM POSSIBLE 41		Total MAXIMUM POSSIBLE 96	
	DECILE	SCORE	%	SCORE	%	SCORE	%	SCORE	%
Group 1B	8th	7	28	17	57	16	39	36	37
	5th	4	16	11	37	10	24	25	26
Group 2	8th	11	44	20	67	21	51	47	49
	5th	6	24	14	47	12	29	32	33

German		Part I Completions MAXIMUM POSSIBLE 25		Part II Definitions MAXIMUM POSSIBLE 25		Part III Anecdotes MAXIMUM POSSIBLE 36		Total MAXIMUM POSSIBLE 86	
	DECILE	SCORE	%	SCORE	%	SCORE	%	SCORE	%
Group 1B	8th	12	48	12	48	16	44	37	43
	5th	6	24	7	28	8	22	19	22
Group 2	8th	13	52	13	52	18	50	39	45
	5th	8	32	7	28	9	25	23	27

Spanish (Form A)		Part I Completions MAXIMUM POSSIBLE 25		Part II Definitions MAXIMUM POSSIBLE 25		Part III Anecdotes MAXIMUM POSSIBLE 38		Total MAXIMUM POSSIBLE 88	
	DECILE	SCORE	%	SCORE	%	SCORE	%	SCORE	%
Group 1B	8th	9	36	14	56	19	50	39	44
	5th	5	20	10	40	13	40	28	32
Group 2	8th	12	48	18	72	23	61	50	57
	5th	6	24	12	48	15	39	33	38

At the upper level range, the groups considered are (1) students who have completed *no* more than two years, (2) students who *have* completed more than two years of college study. The corresponding norm-groups are 11 and 12.

TABLE VII-2

Percentages attained at 8th and 5th deciles

Upper Level

French (Form A)		Part I Definitions MAXIMUM POSSIBLE 25		Part II Anecdotes MAXIMUM POSSIBLE 40		Part III Dialogue MAXIMUM POSSIBLE 15		Total MAXIMUM POSSIBLE 80	
	DECILE	SCORE	%	SCORE	%	SCORE	%	SCORE	%
Group 11	8th	14	56	24	60	6	40	41	51
	5th	8	32	18	45	2	13	27	34
Group 12	8th	20	80	33	82	9	60	59	74
	5th	14	56	23	58	6	40	43	54

German		Part I Definitions MAXIMUM POSSIBLE 25		Part II Anecdotes MAXIMUM POSSIBLE 40		Part III Dialogue MAXIMUM POSSIBLE 15		Total MAXIMUM POSSIBLE 80	
	DECILE	SCORE	%	SCORE	%	SCORE	%	SCORE	%
Group 11-12 (Combined)	8th	10	40	23	58	9	60	40	50
	5th	6	24	13	32	6	40	22	28

Spanish (Form A)		Part I Definitions MAXIMUM POSSIBLE 25		Part II Anecdotes MAXIMUM POSSIBLE 45		Part III Dialogue MAXIMUM POSSIBLE 15		Total MAXIMUM POSSIBLE 85	
	DECILE	SCORE	%	SCORE	%	SCORE	%	SCORE	%
Group 11	8th	13	52	28	62	8	53	46	54
	5th	8	32	21	47	4	27	34	40
Group 12	8th	16	64	31	69	7	47	52	61
	5th	12	48	25	56	4	27	38	45

With these reservations in mind, we may observe at the lower level a generally weaker performance in French than in the other languages. This is apparent even in the totals at the eighth decile, though at the fifth decile only examination of the part-scores will reveal the French as essentially weaker than the German. That is to say, the French scores are disproportionately high on Part II, the section measuring Phonetic Accuracy (which appears only in the French test). This part requires the matching of very short sequences of phonemes with one of four printed units. Correct response thus does not demonstrate the kind of aural comprehension which is measured by completions, definitions or anecdotes, that is to say, both the apprehension and the interpretation of utterances of several words. Since French orthography is extremely inconsistent with the spoken form, some training in the matching of sounds and letters is necessarily given to students of French even in courses not aiming at any high development of aural skill. Most students can distinguish short sequences of sounds (especially when these sounds need not be interpreted, i.e., related to some meaning), and their ability to do so is revealed among the Part II scores even of those who have little ability to comprehend sustained utterances. All students of a language work largely with its written form; and in French, because of the differences between the written and spoken forms, students doubtless need longer and more rigorous training in aural comprehension than do learners of other languages such as German and Spanish. The upper-level French percentages show, moreover, that in the long run students do learn to understand spoken French as well as spoken German or Spanish; and if we are to give weight to the percentages attained by French Group 12, it would appear that advanced students of this language even surpass students of other languages in aural ability. It may be that once the initial phonetic difficulty has been hurdled in French, its basic lexical and syntactic patterns make it an easier language than some others for English speakers to understand.

In German, the lower-level scores are higher on Part I than in Spanish, but lower on the other parts of the total; the upper-level scores are the lowest of the three languages except on Part III, the Dialogue. The German samples tend to be small, and it is probably not safe to conclude that there is significance in these particular differences. However, a partial explanation may lie in a difference of style between colloquial and formal (written) idiom which is greater in German than in French or Spanish. As we mentioned earlier, the anecdotes in the German tests are bookish in character, and may have presented patterns of style which more than ordinarily baffled the listeners. In confirmation of this hypothesis the percentages on the upper-level German dialogue, which is considerably more colloquial in its phrasing than the anecdotes, are higher than any other Part III percentage except those of French Group 12.

In Spanish the percentages are consistently higher than in the other languages in all but the upper-level advanced group, where they fall short of the corresponding French figures. This pattern suggests the hypothesis that students of Spanish, aided by a reliable correspondence between its spoken and written forms, develop a minimum of aural control relatively early but do not thereafter increase this control proportionately with students of French.

Aural performance in Russian cannot be fairly compared with that in the three more widely taught languages. The percentages of achievement for the total tests in Russian are as follows: Lower Level - 8th decile, 76%; 5th decile, 43%; Upper Level- 8th decile, 74%; 5th decile, 57%. Several factors make the situation different in Russian. Classes are small and student motivation is high; virtually no one takes Russian merely in order to fulfill a language requirement. Those who elect Russian have their own reasons for doing so, usually related to their future careers or their cultural interests or backgrounds. Many of them are of Russian, Ukrainian or Polish extraction and have heard, if not actually spoken, in their homes one or another of the closely related Slavic tongues. Those who have not this advantage are apt already to have studied one or more foreign languages, and they come to Russian with a proved interest in, and aptitude for, foreign-language study. Russian is popularly believed to be an extremely difficult language for speakers of English to master, and few students attempt it who do not know in advance that their chances of success are better than average.

The predominantly low ranges of scores, in all the languages at both levels, demonstrate clearly enough that the average American student of a second language is not equipped at the end of one, two, or even three years of college study to understand adequately a native speaker transmitting unfamiliar material through a mechanical medium. Combining all three languages and both levels, we find that at the 8th decile the percentages range from a high of 74% to a low of 37%, and that at the 5th decile (the median) they range from a high of 54% to a low of 22%. Furthermore, a glance back at Table III (page 72) shows us that the average student in a current experimental course — a course in which he is trained through special methods to handle the spoken form of the language, and can allegedly comprehend with complete facility familiar sequences of speech uttered by speakers who are also his teachers — is no more successful than the conventionally taught language student when confronted with an instrument designed to measure his aural ability in absolute terms.

Let us attempt to examine the factors contributing to the equal insuccess of conventionally taught and aurally trained students. In the case of conventionally taught students, the basic cause is quite simple: their aural training has been, at worst, non-existent, and at best casual or inadequate. In many so-called reading-courses there is frankly no other objective than ability to interpret the printed page; in others there is "oral-aural practice" in class, which probably consists mainly of oral reading and contributes little to the development of aural skill, especially if the students hear each other's accents more than the teacher's. In the so-called "all-purpose" or "eclectic" course, where aural ability is claimed as one of the objectives, the most skillful and well-meaning teacher has all too little time to devote to the painfully slow process of training his students' ears. Furthermore if this teacher is not a native speaker of the language, even if he devotes a maximum of time and energy to aural training, the students are apt to end up understanding him but no one else. Before being exposed to the Investigation's aural tests, many students have never heard a foreign language reproduced mechanically; they feel (and are) at a disadvantage when faced with a phonograph record. But if they have never heard native speakers, they would not perform with notably greater success if the person who made the recordings were to step before them and deliver the test himself, with gestures.

The students aurally trained in current experimental courses obviously enjoy advantages which the other students lack. They hear the language daily in their drill sessions, and are trained to understand native speakers at a regular conversational speed. Yet these advantages are offset by another factor, the ultimate effect of which is to reduce them to the same level of non-comprehension (so far as the tests are concerned) as the untrained group. The topics discussed in their conversation practice — menus, markets, and motors — are not the subjects treated in the definitions and anecdotes of the aural tests, and the vocabularies are to no small extent different.[3] Thus their scores could hardly have been perfect, even if they had understood every word – and every *sequence of words* – which they had met and learned in their oral-aural work. It is evident, however, that regardless of how many known words they actually *apprehended,* they did not *comprehend* the total flow of speech as a whole.[4] That is, instead of "hanging on" and filling in from the context and from inference the gaps caused by unfamiliar words, they missed the small structural words (those which give precise form to an utterance), were disconcerted by the occurrence of unknown words, and so lost the continuity of the sentences and anecdotes. And it is safe to assume that they failed of over-all comprehension because, despite their special training, they were still far short of having acquired near-native skill in aural comprehension.

It would have been interesting to learn exactly how high a level of performance would be reached by a representative sample of persons having complete native control of the language tested. Such a group would not, of course, all achieve perfect scores for at least two reasons: (1) failure to remain completely attentive throughout the test, and (2) failure to cope unerringly with the mechanics of the test and to handle the *English* answer choices within the time-limits imposed. In any case, no such group of bilinguals was accessible. We did, however, have the opportunity to test several hundred Spanish students in the southwestern United States who came from Spanish-language backgrounds. Both college and high-school groups were tested at each level. Scores and percentages obtained were as follows:

---

[3] Cf. the range of overlap cited in Chapter III, p. 47.

[4] If they had, we estimate that the incidence of unknown words would have held their possible maximum performance to roughly 80%. We sought data on which to base this estimate by giving the Spanish tests (both levels, Form A) to a person who knew no Spanish whatever, translating into English as we read the material all the words occurring in Treviño's *Spoken Spanish*. The resulting total scores were 66 on the lower-level (75%) and 70 on the upper-level test (82%).

	8th decile	*5th decile*	*Mean*
Lower Level - College Group (N=44)	82 (93%)	74 (84%)	73.9
Lower Level - High-School Group (N=267)	62 (70%)	50 (57%)	47.8
Upper Level - College Group (N=62)	74 (87%)	68 (70%)	65.2
Lower Level - High-School Group (N=233)	53 (62%)	46 (54%)	42.7

None of these groups could be considered homogeneous in competence. The college groups were more so than the high-school groups, and their greater success on the tests, at either level, seems due in part to this factor. All were about equally competent in English; some spoke Spanish in homes at least moderately cultivated; others, while not fluent themselves, associated frequently with native speakers and understood the language with considerable facility. It is noteworthy that the lower-level group was more successful than the upper-level group, a fact which at once suggests that students of this type command a somewhat limited vocabulary.

The high-school students tested were classified within the school system as "Spanish-speaking." We may suspect that this distinction is not entirely linguistic, however, and that ethnic and social factors play some part. Some of the students, while they had Spanish names, performed as though they knew very little Spanish. It was their scores which brought the group means below the medians. Many of them doubtless came from families who no longer used much Spanish in the home, at least with the children; others of them probably had limited control of English. The majority of them, moreover, belonged to a social group having relatively little breadth of culture; this meant that their Spanish was probably limited to a vocabulary of a few hundred words, not greatly expanded since early childhood. Even on the lower-level test, half of them understood only 57% or less of the material.

Even though the southwestern students tested in Spanish could not qualify as native speakers of the language, they assumedly possessed greater aural skill than the typical American student. In the light of this assumption, their superior performance on the aural tests indicates a positive correlation between the ability to understand Spanish and success on these tests, however imperfect they may be. And though reflecting two rather distinct levels of aural skill, the two "Spanish-speaking" groups tested both stand between two extremes: that of the fully competent native speaker on the one hand, and the average American student of a second language on the other.

ORAL PRODUCTION

While the country-wide response to publicity regarding our aural comprehension series was considerable, there was comparatively little demand for the oral production tests. For one thing they were not completed and available until late in the academic year 1945-1946 – approximately a year after the aural tests went into use. This lateness combined unfortunately with the great difficulties involved for any teaching staff in setting up a schedule for the administration of the tests.

Since the tests require fifteen to twenty minutes for each student, it was out of the question to attempt any large-scale use of them. Even in the centers collaborating especially with us to furnish large bodies of assorted data, it was impossible to develop a concerted plan. The most that we acquired was a sampling, which was haphazard even in the case where we ourselves spent two days on a campus doing the testing. A further disturbing factor was the propensity of teachers to select as samples, when full groups could not be tested, students whose ability to speak the language was better than average for their group. Also, in one college 27 out of the 56 students tested at the upper level were advanced students, and some of those 27 were language majors taking junior or senior courses not included at all in our aural and visual testing program at that college. Again, 16 intermediate students at that same college were given the lower-level instead of the upper-level test. Thus the ranges of scores obtained would undoubtedly be lower than they are if the student samplings had been more representative of poor, average and good students and had been less heavily weighted with advanced students.

The following table shows the extent of the oral testing program, which consisted of samplings of both conventional and experimental students.

	LOWER LEVEL		**UPPER LEVEL**	
	Number of Cases	*Number of Institutions*	*Number of Cases*	*Number of Institutions*
French	23	2	51	3
German	27	4	33	4
Spanish	43	4	26	3

For obvious reasons we made no attempt to standardize the oral tests. Table VIII shows the mean scores obtained for the separate parts on each test, with percentages given in parenthesis after each score to indicate the actual level of performance.

At the lower level the scores on all three parts of the test are lowest in French and highest in Spanish. Although, as we shall see in a moment, all the scores are of somewhat dubious validity, the contrasts among the scores for the three languages at the lower level may possibly be regarded as indicative of relative degrees of difficulty inherent in the earliest oral efforts in these three languages. Problems of pronunciation certainly loom largest in French, and are

TABLE VIII

Oral Production Tests: Mean Scores (Raw) for All Groups Combined

	Part I Picture Series	Part II Sustained Speech	Part III Conversation	
	MAXIMUM POSSIBLE 40	MAXIMUM POSSIBLE 8	MEAN	MAX. POS.
French				
Lower Level (N=23)	12.3 (31%)	3.3 (41%)	24.1 (30%)	80
Upper Level (N=51)	18.7 (47%)	4.8 (60%)	36.7 (46%)	70
German				
Lower Level (N=27)	14.5 (36%)	4.2 (53%)	35.4 (37%)	96
Upper Level (N=33)	21.9 (55%)	6.4 (80%)	51.9 (58%)	90
Spanish				
Lower Level (N=43)	25.5 (64%)	5.1 (64%)	35.5 (51%)	70
Upper Level (N=26)	23.7 (59%)	5.1 (64%)	34.2 (50%)	68

perhaps greater than those of inflection and word-order in (simple) German; Spanish is relatively easier, both phonetically and syntactically, in the earlier stages. The contrasts among scores from language to language are not as marked or as consistent at the upper level, thus perhaps reflecting the existence of lesser difficulty differentials — though not of decreasing difficulty — in the intermediate and advanced stages of oral development.

It will further be noted that the performances on Parts I and III are very close together in both levels of French and German, while in Spanish Part I runs higher than Part III. This may mean simply that in French and German the parts were of about equal difficulty but that in Spanish, Part I was easier. In Parts I and III samples of acceptable response were furnished for each item on the rating sheet, and with these guides most administrants were able to judge the responses item by item with passable reliability. On Part II, however, useful models for the production of sustained speech could not efficiently be supplied; thus, despite the rating scale supplied, examiners necessarily relied on the standards of excellence they were accustomed to set for their own students with respect to fluency, vocabulary, pronunciation and grammar. It is rather unlikely that a given student's production would vary greatly in quality from one part of the test to another, and especially unlikely that he would perform more expertly on the one section in which he had to formulate his own ideas as he talked; hence we are inclined to see in the relatively high Part II scores a failure of the examiners to rate as objectively as on the other parts.

We ourselves administered the oral tests in one college and in one high school, while in a second college the students' production was recorded on discs and sent to us for rating. The scores we assigned tended to be lower than those given by local administrants, and were consistently so on Part II. Below are transcripts of the production of three students who took the lower-level Spanish test in College A, and of three who took the upper-level German in College J.[5]

Samples of Spanish Oral Production

The Spanish students were completing a one-year beginning course which met five hours per week. This course was of the conventional type, in which nevertheless there was an attempt to develop oral and aural skills as much as possible. The responses on the oral test were recorded on discs, and scored by the Investigation staff. The students are designated respectively as A, B, and C. Their responses appear in the right-hand column opposite those suggested as a standard in the rating-sheet for examiners.

[5] The programs of these colleges are described in Chapters V and VI.

PART I

Picture Series: 40 points
A's score: 21
B's score: 17
C's score: 11

Standard Response	*Recorded Responses*
1. Los niños están jugando al tenis.	A: Los ninõs están jugando. B: Los muchachos está jugando el tenís. C: Los dos chicos juguen el tenís.
2. La mujer está comiendo fruta.	A: La muchacha.. (ah)..comiendo .. (no).. come fruta. B: La muchacha come .. (ah)... C: La muchacha comé la fruta.
3. El muchacho está vendiendo periódicos.	A: El muchacho vende los periódicos. B: El muchacho venden el diario. C: El chico vendé el periádico.
4. La criada está lavando los platos.	A: La criada lleva las platas. B: La criada llavan los plet/... platos. C: La muchacha lleva la ...
5. La gente va a la iglesia.	A: La familia van ... va a la iglesia. B: La gente van a la iglesia. C: La gente van a la inglés.
6. La familia está escuchando el radio.	A: La familia escuchando ... (no) ... escuchan a la radio. B: La familia echan el radio. C: La familia escuchan el radio.
7. El joven se despierta.	A: El muchacho despiert/...se despierta. B: El muchacho se despierta. C: El chico se desperté.

8. Los niños salen de la escuela.	A: Los niños andan. B: Los muchachos van a la escuela. C: Los alumnos salen a la escuela.
9. El hombre corre hacia el autobús.	A: La mucha/...(or) El muchacho está coriendo a la bus. B: El muchacho ... C: El chico está corriendo a coche.
10. El perro tiene un zapato en la boca.	A: El pero tiene un [z]ápato[6] en el ... en la boca. B: El perrito lleva la zapato. C: El peró tiene un [š]rape.
11. La chica está cogiendo flores.	A: La muchacho coge los flores. B: La muchacha coge los flores. C: La muchacha es ... picar las flores.
12. Los soldados están marchando.	A: Los soldados son marchando. B: Los soldados marcha. C: Los soldados están marchando.
13. El artista está pintando un cuadro.	A: El hombre pinta. B: El artista pinta el retrato. C: El chico pinta el carta.
14. La maestra está escribiendo en el tablero.	A: La profesora escribe en la pizarra. B: La profesora escriba la lección. C: La profesora escribe la lección en el pi[z]ero.
15. La taza está rota.	A: La disca es rota. B: La taza echán per/... pedazos. C: ...
16. La chica baja la escalera.	A: La muchacha va a la baja. B: La muchacha ... C: La muchacha ... es ...s/...

---

[6] Letters in brackets are phonetic representations of sounds not occurring in Spanish.

17. El hombre está encendiendo su cigarro.	A: El caballero fueg/...fuegue el cigaro.
	B: El caballero está fumando.
	C: El muchacho fiuma cigarillo.
18. El señor se pone el abrigo.	A: El hombre mete la ábrigo ... el ábrigo.
	B: El señor lleva el abrigo.
	C: El pobre es se vestí.
19. El niño se acuesta.	A: La niña se ... ac/...acoste.
	B: La muchacha ...
	C: El muchacho se acuesta.
20. El cartero está llamando en la puerta.	A: El hombre de la c/. . . las cartas llama a la puerta.
	B: El hombre llama a la puerta.
	C: El muchacho llama a la puerta.

PART II

Sustained Speech: 8 points

A's score: Fluency 1, Vocabulary 1, Pronunciation 2, Grammar 1.
B's score: Fluency 0, Vocabulary 1, Pronunciation 0, Grammar 1.
C's score: Fluency 0, Vocabulary 1, Pronunciation 0, Grammar 0.

Stimulus: "You are talking with a Spanish-speaking person who has never been to the United States. Describe to him the town or city in which you live." *Second stimulus* (if necessary): "This person is also interested in what a North American home looks like. Describe to him the home in which you (or your parents) live."

A: "[Place-name] es una...(ah)...ciudad muy hermosa. Tiene un popiulación de trescientos mil. Es una ciudad sichuada en muchas montañas (*laughs*) y es... (ah)... y tiene una río: el río que coriendo en la... en medio de la ciudad. Tiene una catedral que es muy hermosa y es muy famosa en la páis. Es la capital de la... es/.. de... del Estado ________. Los edificios son... no es muy largos; es... tiene... tienen diez, y cuatorce, y más quince pisos. Las calles coriendo..." [*Next stimulus*] "Mi casa... tiene siete...(ah)... habitaciones. Es... tiene dos pisos y es de...(ah)... madera. Es amarillo, y

se es/...tá sichuada en e/... encima de un... encima de un gil. Tiene muchas ... (ah)... ventanas que son muy largas y muy bonitas."

B: "El pueblo donde vivo está muy pequeño. La calles as/...está larga y estrecha. Las casas astá antigua. Las calles tienen..." [*Next stimulus*] "La casa donde vivo astá pequeño. La casa está rodeado por un jardín. El jar... En el jardín hay árboles, flores; flores son de colores de rojo, azul, amorilla, y...(ah)...esta primavera los árbol/...las árboles astá... astán verdes. La casa...(ah)...tiene...(ah)... seis cuadros: la cócina, la sala, tres... álcobas..., y un cuadro de baño. En la sala está dos sillóns, un piano, un gran alfambra..."

C: "[Place-name] está sichuado en el lago de Lake__________. Tiene papiule[š]ión mil quinientos vente, y es un pobla[š]ión muy hermosa. En el pobla[š]ión hay muchas cosas hacer. Estoy ya no viviendo en__________. Mi madre y mi padre están viviendo allí. Hay muchos barcos en el lago." [*Next stimulus*] "Mi madre y mi padre viven en una casa sichuada al...en el calle de Oak. La casa no es muy grande pero sirve...*les* sirve muy bien. La casa tiene dos pisos y cuatro cuadros en cada piso."

PART III

Conversation: 70 points

A's score: 43
B's score: 27
C's score: 13

V = Voice on record ***R = Standard response***	***Recorded Responses***
V: ¿Bueno, cómo le va?	
R: Estoy cansado y me duele un poco la cabeza.	**A: Muy bien...(oh)... Yo estoy cansado. Tiene una pena de...(ah) ...cabello.**
	B: Está cansado y tengo un dolor de cabeza.
	C: Muy cansado y dolor de cabe[z]a.

V: ¿Por qué, hombre?	
R: Acabo de volver de un largo viaje.	**A: Acabo de volver desde una viaje largo.** **B: Acabo de llegar de un largo viaje.** **C: Vol/... Vol/...**
V: Usted debe descansar por algunos días.	
R: Me gustaría mucho, pero tengo que volver al trabajo mañana.	**A: Quiero que podré, pero debe estar ...estaré de vuelta mañana.** **B: (No sé.)** **C: Quiere que pued/...puedo, pero...**
V ¿Sigue usted todavia en el banco?	
R: Hace tres años que dejé ese puesto.	**A: Tres...Hace tres años que salí de aquel posición.** **B: (Oh, I quit right now!)** **C: He salido...**
V: ¿Que hace usted ahora?	
R: Ahora estoy enseñando inglés, y me gusta mucho más esta clase de trabajo.	**A: Ahora enseño inglés, y me gusta esta trabajo much mejor.** **B: ....** **C: Ahora estoy el inglés, y...**
V: Muy bien. A propósito, ¿sabe usted que ya tengo dos niños?	
R: Le felicito. ¿Cuántos años tienen?	**A: ¿Cuántos años tiene?** **B: ¿Cuántos años tiene?** **C: Congrachula[š]ión. ¿Tiene...**
V: La chica tiene dos años, y el nene seis meses.	
R: ¿Quiere usted almorzar conmigo?	**A: ¿Acompañéme al comedor?** **B: ¿Quiere usted va...** **C: ¿Vino...**

V: Cómo no? Con mucho gusto. ¿A dónde iremos?	
R: Conozco un restaurante en donde sirven comidas estupendas.	A: Conozco uno café que sirve comidas muy buenas. B: Sé un restaurante que sirven las comidas buenas. C: Conocé un café que sirve comedor muy bien.
V: Pues muy bien. Vamos.	
R: Tengo un coche abajo.	A: Tiene un automovíl debajo. (¡Tengo!) B: Tengo un automóvil arrib ... abajo. C: Tengo un coche escaldas abajo.
V: ¿Hace mucho tiempo que tiene usted coche?	
R: Aprendí a conducir el año pasado, pero este coche es de mi padre.	A: Aprendí a manejar el año pasado, pero esta automovíl pertenecer... pertenecé al mi padre. B: Esta automóvil está el de mi padre. C: Canduzco el año pasado, pero este coche...
V: ¿Su padre vive todavía en el campo?	
R: No. Él y mi madre se quedan en la ciudad hasta el verano próximo.	A: Mi padre y mi madre están quedando en la ciudad hacia la ... B: Mis padres viven en la ciudad hasta el verano que viene. C: Mi madre y yo ...
V: Recuerdo que sus padres tenían una casa muy hermosa.	
R: Desgraciadamente quieren venderla.	A: Quieren que ven ...vende...lo vende. B: Quieren que... que la venden. C: Quieren vender el coche.

V: Bueno, ya estamos.	
R: ¿Quiere usted que pida algún vino blanco?	A: ¿Debo le man/... de mandar algún vino blanco? B: .... C: ¿Quería que...
V: Sí, muchas gracias.— Bueno, ha sido un almuerzo estupendo. ¿Quiere usted ir al cine ahora?	
R: Gracias, no puedo. Quiero volver temprano a casa.	A: No, gracias. Quiero que volver a casa temprano. B: No, gracias. Tengo estoy en casa muy temprano. C: No, gra[š]ias. Quiero a casa temprano.
V: Ah, ¿sí? ¿Por qué?	
R: Estoy esperando un telegrama importante esta tarde	A: Espero una telegrame importante esta tarde. B: Espero un telegrama importante esta tarde. C: Esperó una telegrám muy importante.

Sample of German Oral Production.

The German students were completing an accelerated second-year college course at the rate of fourteen hours per week through one summer (1946). Considerable attention was paid in this course to the development of oral and aural skills. The responses were transcribed by a person present, without any attempt to reproduce faulty pronunciation or hesitations. Scores by items were not made available to us. The three students are designated respectively as A, B, and C.

PART I

Picture Series: 40 points

A's score: 12
B's score: 23
C's score: 20

Standard Response	*Actual Response*
1. Die Frau strickt einen Strumpf.	A: Eine Frau... B: Eine Frau macht einen Socke. C: Die Dame macht...
2. Der Jäger läuft vor dem Löwen davon.	A: Ein grosser Tier fällt nach einem Mann. B: Der Löwe will den Mann essen. C: Der Löwe verfolgt den Mann.
3. Der Skiläufer macht einen Sprung.	A: ... B: Der Mann lauft mit Ski. C: Der Mann macht Skilauft.
4. Der Polizist regelt den Verkehr.	A: Der Schutzmann ampelt das Verkehr. B: Der Verkehrschutzmann regelt der Verkehr. C: Der Polizeimann hält die Auto.
5. Der Fleischer wiegt das Fleisch ab.	A: Da ist ein Mann, der Fleisch verkauft. B: Der Mann wäscht den Fleisch. C: Der Mann... Der Fleischer...
6. Das Mädchen winkt dem Schiff nach.	A: Ein Mädchen grüsst in Dampfer. B: Die Frau sagt auf Wiedersehen zu dem Schiff. C: Die Dame sagt auf Wiedersehen.
7. Der Bauer pflügt sein Feld.	A: Ein Bauer arbeitet in der Feld. B: Der Mann arbeit in dem Feld. C: Der Bauer arbeitet im Feld.
8. Das Dienstmädchen klopft den Teppich aus.	A: Sie schlägt den Teppich. B: Die Frau schlägt den Teppich. C: Das Dienstmädchen schlägt die Wolle.

9. Der Kellner weist der Frau einen Platz an.

A: Eine Frau geht ins Hotel, und der Kellner gibt ihr einen Platz.
B: Der Ober geben die Frau einen Platz.
C: Der Ober sieht die Dame...der Dame einen Stuhl.

10. Er giesst ihr Wasser ein.

A: Er giesst Wasser ins Glas.
B: Der Ober schenkt den Wasser ein.
C: Er giesst fliessendes Wasser in einem Glas.

11. Er zeigt ihr die Speisekarte.

A: Die Dame sieht das Menü.
B: Der Ober gibt die Frau eine Speisekarte.
C: Er gibt ihr die Speisekarte.

12. Er notiert die Bestellung.

A: Sie gibt dem Kellner, was sie essen will.
B: Die Frau bestellen, was sie zu essen wünscht.
C: Er nimmt ihren Berichte.

13. Der Räuber macht einer Überfall.

A: Ein Mann ist von einem Dieb...
B: Man sieht einen Mann attacken.
C: Ein Dieb hält einen Mann auf.

14. Er nimmt dem Mann sein Uhr weg.

A: Er nimmt eine Taschenuhr aus.
B: Er nimmt seine Uhr.
C: Er nimmt seine Uhr.

15. Der Polizist schlägt den Räuber auf den Kopf.

A: Und dann kommt der Schutzmann und schlägt den Dieb.
B: Der Schutzmann schlägt den Kerl auf den Kopf.
C: Ein Schupo schlägt auf den Kopf.

16. Der Verbrecher ist im Gefängnis.

A: Da ist er verschlossen.
B: Der Kerl ist gehält.
C: Der Dieb geht ins...

17. Zwei Autos sind ineinander gefahren.	A: Wir haben hier ein Unfall, zwei Autos kommen zusammen. B: Zwei Autos haben einen Zusammenschluss. C: Zwei Autos fahren gegeneinander.
18, Der Verunglückte wird auf der Tragbahre davongetragen.	A: Ein Mann ist verletzt und man ruft der Krankenauto. B: Zwei Männer tragen einen Mann auf einen Krankenträger. C: Die Krankenträger tragen einen Mann zum Krankenwagen.
19. Er wird in den Krankenwagen gelegt.	A: Der Mann... B: Sie stellen den Mann in den Krankenwagen. C: Sie stecken in den Krankenwagen.
20. Die kaputten Autos werden abgeschleppt.	A: Dann kommt der Lastauto. B: Zwei Lastautos sie ziehen die Autos von der Zusammenschloss. C. Zwei Lastautos nehmen die beide Auto her.

PART II

Sustained Speech: 8 points

A's score: 5
B's score: 5
C's score: 5

Stimulus: "You have found a sympathetic listener and are about to launch upon the story of your life. Start as early in your life as you wish and go on until the examiner stops you."

A: "Es ist nicht in Ordnung von mich zu sagen. Ich war in 1927 geboren. Ich wohne in einen Ort mit... und ich habe eine grosse Familie, und ich habe fünf Schwestern und zwei Brüder, und bin [place-name] seit zwei Jahr, und ich hoffe einen Tag ein Chemikin zu werden. Die Sommer habe ich sehr gern. Ich arbeite in einem Laboratorium. Viele Menschen glaubt, dass ein Mädchen kann nicht... solte Chemie nicht studieren. Aber ich...es interessiert mich viel."

B: "Ich bin 1923 geboren in [place-name]. Ich habe dort gelebt, und ich bin dort jetzt lebend. Meine Mutter war dort verheiratet. Mein Vater hat eine Fabrik, wo er artificial Leder macht. Wir haben ein Haus, in wo wir im Sommer wohnen. Ich bin 1945 nach [place-name] gekommen, und ich habe hier bis 1943 geblieben. 1943 bin ich ins Armee gegangen. In das Armee ich war im Luftwaffe, und dort war ich ein Officier der Gunnery. In 1945 bin ich in den Pacific gegangen. Dort bin ich über Japan gefliegen. Im Juni 1945 wir haben einen Zusammenschloss gehabt, und einen Monat dahinter habe ich nach Amerika in einem Krankenairplane geflogen."

Stimulus: "You have met a young German in Europe who seems to you to have the makings of an outstanding American citizen. You resolve to try to convince him that he should emigrate to the United States. Talk to him about the United States, so that you may help him to decide whether he would like to come."

C: "Sie sind wie ein guter Mann aus. Sie wollten einen guten amerikanischen Mann sein. In Amerika haben wir viel Freiheit. Wir können viel tun das Sie in Deutschland können nicht. Der President ist bei die Leute chosen und da er ist für vier Jahre der Führer. Man muss nicht....."

PART III

Conversation: 90 points

A's score: 52
B's score: 47
C's score: 49

V = Voice on record *R = Standard Response*	*Actual Response*
R: Haben Sie fur dieses Wochende etwas vor?	A: Geht es über die Wochenende? B: ... C: Was wollen Sie Wochenende machen?
V: Nein, ich habe mir garnichts vorgenommen.	
R: Wie wäre es, wenn Sie und Ihre Frau einen Ausflug mit uns zusammenmachen?	A: Warum gehen Sie und Ihre Frau.... B: Warum kommen Sie mit Ihre Frau zu... C: Kommen Sie doch mit mir auf ein Picnic.

V: Das ist eine ausgezeichnete Idee. Was haben Sie vor?	
R: Meine Frau hat an eine Autofahrt aufs Land gedacht.	A: Meine Frau sagt, dass wir ins Ausland mit ihrem Auto gehen. B: Meine Wife will ins Land in das Auto fahren. C: Meine Frau will zum Land mit Auto fahren.
R: Ich weiss einen abgelegen See oben in den Bergen.	A: Ich kenne einen sehr schönen See hoch in Bergen. B: Ich kennen ein See hoch in den Bergen. C: Ich kenne ein herrlicher See in den Bergen.
R: Mir gefällt es dort weil es schön kühl und angenehm ist.	A: Das gefällt mir weil es so kühl und schön ist. B: Ich habe das gern, weil er sehr kühl und ist bequem. C: Es gefällt mir, weil es kühl und angenehm ist.
V: Das passt mir gut. In welchem Auto wollen wir fahren?	
R: Mein Auto ist grösser als Ihrs.	A: Aber mein Auto ist grösser als Ihrer ist. B: Mein Auto ist grösser als Ihres. C: Mein Auto ist grösser als Ihrs.
R: Können Sie Ihren kleinen Neffen mitbringen? Dann hat meine Tochter jemand zum Spielen.	A: Warum bringen Sie nicht seinen kleinen Neffen mit, und ich will bringen meine Tochter. B: Können Sie eine Neffe mitbringen, und meine Tochter wird... C: Können Sie Ihren Neffe bringen; meine Tochter will mit ihm spielen.

V: Natürlich, wenn seine Mutter nichts dagegen hat. Ich werde sie heute Abend fragen. Wir sorgen natürlich fürs Essen.

R: Wenn Sie das Essen mitbringen, dann müssen Sie mich fürs Benzin bezahlen lassen.

A: Wenn Sie das Essen bringen, dann muss ich die Benzin bezahlen.

B: Wenn Sie das Essen bringen, müssen Sie mich um das Benzin zahlen dürfen.

C: Wenn Sie das Essen bringen, dann muss ich für Benzin zahlen.

R: Sie müssen mir versprechen, dass sich Ihre Frau nicht zu viel Umstände dabei macht.

A: Aber Sie müssen mir versprechen, dass Ihre Frau...

B: Versprechen Sie mich, dass Ihre Frau nicht zu viel...

C: Ihre Frau mussen nicht zu schwer arbeiten.

V: Wann wollen wir uns dann treffen?

R: Sonnabend früh punkt sieben Uhr, wenn dass nicht zu früh ist.

A: Wenn es nicht zu früh für Ihnen gehen Sonnabend morgen...

B: Sonnabend früh um sieben Uhr, weil das ist nicht zu früh ist.

C: Samstag Morgen um sieben Uhr, wenn das nicht zu früh ist.

V: O nein, es ist garnicht zu früh. Und wo wollen wir uns treffen?

R: Ich werde Sie bei Ihnen zu Hause abholen.

A: Ich will Sie...Ich will Ihnen zu Hause...

B: Ich treffe mit Ihnen um ... bei Ihnen ...

C: Ich will Ihren ... zu Ihrem Hause.

V: Das passt mir ausgezeichnet. O, bevor ich es vergesse, ein Freund von mir hat mir zwei Karten für die Oper heute abend geschenkt. Meine Frau und ich können leider nicht gehen. Möchten Sie sie gerne haben?

R: Das ist furchtbar nett von Ihnen. Welche Oper wird gegeben?

A: Das ist sehr nett von Ihnen. Welche Opern gibt es?

B: Danke sehr. Welche Oper gegeben ist?

C: Das ist sehr nett von Ihnen. Welche Oper spielt?

V: *Die Meistersinger von Nürnberg.*

R: Ich kann diese Oper immer wieder hören.

A: Das Oper kann ich über und über hören.

B: Ich kann jede Oper immer hören.

C: Ich könnte das wieder und wieder hören.

V: Das freut mich sehr. Hoffentlich gefällt Ihnen die Vorstellung.

R: Ich muss mich jetzt beeilen, sonst verpasse ich meinen Zug.

A: Ich muss eilen, oder ich will meinen Zug verlassen.

B: Ich muss mich eilen, oder meine Zug verpassen werde.

C: Ich musse eilig gehen denn ich mein Zug nicht...

V: Na, dann will ich Sie nicht länger aufhalten.

R: Grüssen Sie Ihre Frau von mir, und rufen Sie mich doch morgen nach dem Abendessen an.

A: Grüssen Sie meine Frau für mich, und rufen Sie mich morgen Nacht.

B: Grüss Ihre Frau, und reden Sie mich nach das Abendessen.

C: Rufen Sie mir morgen an.

Evaluation of Oral Performance

We pointed out at the beginning of this section on oral production that our test results are based on small samples, that those samples are unrepresentative in that they are unduly loaded with advanced students and with top performers, and that the rating of achievement by teachers themselves was unavoidably subjective. Even though the resulting figures must be regarded as high by comparison with what would have been obtained from a more representative group rated more objectively, these existing figures are low in terms of total possible achievement. They range from 30% to 64% at the lower level and from 46% to 59% at the upper (exclusive of the unreliably rated Part II). They thus give clear enough evidence that the American student of a second language shows no highly developed ability to speak that language when confronted with a test designed to measure that ability within a vocabulary range not exceeding 3,000 words. This is of course to be expected in the case of students not specifically trained in oral production. The oral tests, however, produced no notably different results with the handful of students in experimental courses giving specific oral training. Such students can undoubtedly outtalk the conventional student in such situations (real or simulated) as form the basis of his particular textbook — he can go to the theater, or lay in supplies, or play games in the language if his book does, because he has learned the specific words, and sequences of words, to use in the given situation. But take him out of his special situation — ask him, for example, to chat with a friend on some other topic, or to make small-talk — and he is only too likely to be tongue-tied. If he knew the words and patterns to use he would doubtless pronounce them more authentically than the conventional student, for he has learned to articulate through direct imitation of a native speaker. But his vocabulary has enormous gaps, and what is more he has not yet developed a flexibility in handling and varying the structural patterns of the language. Thus he is comparatively immobilized when confronted with the unfamiliar test situation. And those who would claim for this student an approximation to native speech must surely realize that on these same oral tests a native speaker would (provided he knew enough about English to follow the directions for responding in Part III), automatically make a perfect score.

SUMMARY

In the present chapter on the Investigation's Aural and Oral Testing Program we have described our apparatus for standardizing the aural comprehension tests, have presented the norms which resulted, and on the basis of those norms have attempted certain comparisons between different college and high-school groups in each language, as well as a few comparisons (subject to reservations) between similar groups in different languages. We have discussed the problems involved in the administration and rating of the oral production tests, and pointed out the impossibility of standardization under the circumstances which prevailed.

On the basis of the results obtained in the two series of tests we conclude that neither the conventionally taught nor the specially trained students were able to understand or to speak the language with near-native competence after studying it up to three years. Further, the student in an experimental course emphasizing the development of aural-oral skills demonstrates no greater proficiency than the untrained student probably because (1) while he admittedly has good comprehension of native speakers and authentic production within limits, those limits are narrow, determined by the topically specialized materials he studies; (2) he has not developed skill in quickly comprehending sustained segments of speech heard for the first time or in easily producing variations of the structural patterns he has learned.

Chapter V

Conventional Groups Not Stressing Aural-Oral Objectives

The norms presented in the preceding chapter constitute one scale for measuring the work of the experimental groups, for those tables give a fairly representative portrayal of the general level of students' aural comprehension as it is defined and measured by our tests. That standard is valuable as far as it goes, but still more exact comparisons are possible if comprehensive testing results from experimental programs can be matched with equally detailed measurement of conventionally taught groups. Seven colleges and one high school did this rather extensive testing for us, administering both the reading and aural tests to classes not taught by any experimental method.[1] In addition, many of them supplied supplementary data about the students (intelligence scores, high-school records, college marks, etc.) parallel to those furnished us by most of the experimental programs. Some of these institutions were also offering experimental courses; but in their case these intramural comparisons are neither so exact nor so revealing as might be expected. Consequently we sought information from other schools as well. Though these conventional groups are mostly at other institutions than the experimental ones, they are at somewhat similar schools—in so far as the diversity of American educational institutions makes for even general similarity. They thus permit good, though not perfect, comparison.

---

[1] The current interest in the aural-oral skills and in intensive instruction may have led some of these institutions to pay somewhat more attention to these abilities and to increase slightly the number of weekly class hours to four or five; but no acoustical laboratories have been installed, no native speakers imported, and no new textbooks adopted.

Furthermore, the study of these conventional classes, quite apart from its value in assessing the work of the experimental courses, is significant in its own right as a picture of current language teaching in this country, its procedures and achievements.

College A[2]

A professor of Spanish at College A (a small liberal arts college) used the Investigation's aural tests in May, 1946. This experience led him to comment to us with such interest and penetration on the basic problems of second-language teaching that we asked him to arrange a full testing program, covering three languages, at College A for the academic year 1946-1947. Through our colleague's untiring collaboration and meticulous care for details, these plans were executed with more than usual efficiency. We present the resulting data as fairly typical of second-language teaching in small liberal-arts colleges at the present moment.

---

[2] Instead of naming the various centers, we shall designate them as College A, College B, High School A, etc. Our aim is not primarily to disguise the identity of the various institutions; readers familiar with current developments in the field will probably recognize most of them, for all other relevant information is reported accurately and in detail. Rather we have in mind two purposes. One is a matter of relative emphasis: in an objective investigation, the identity of a given institution is not a relevant factor in evaluation; and we hope that anonymity will keep the reader continually aware of the most important factor, namely the relation between procedures and outcomes in a program, regardless of where it may be or who may be directing it. Secondly, we wish to avoid all implications of advertising on the part of the institutions involved, and of either praise or censure (as the case may be) on our part. In any event, from the point of view of language teaching in general, the important result is not the success or failure of a particular program at a particular institution but the general knowledge of principles and procedures which can be drawn from these undertakings.

Spanish

Our collaborator at College A describes the first and second-year Spanish courses as follows:

> You probably noticed that our students gave a wide range of performance on your tests.[3] I think it is at least partly due to the fact that we try to carry students of a wide range of ability, with a method sufficiently eclectic for them to progress most in the skills most suited to them.
>
> These students were the first ones to go through a whole year of a system which we have been developing to give a considerable amount of oral work without sacrificing much, if any, of the other aims. It is quite simple, and it may be that no part of it is new, but I haven't found its elements elsewhere. It consists of questions asked about whatever text material we are using, whether the Spanish section of a grammar lesson or a section of a reader. These questions break the sentences of the text into smaller units, and can be answered directly from the text, in some cases requiring a shift of person in the verb or a possessive adjective, in other cases requiring no shift at all. We prepared mimeographed sheets for Keniston's *Learning Spanish;* some of the lessons will be revamped before fall to give more practice answering in the first person. Because the answers can be taken so easily from the text, the students soon get so they almost always give the answer correctly. This means that they are writing and saying correct Spanish, and not forming bad habits. We use the questions and answers in two ways: (1) as ordinary question-and-answer material, the instructor asking the questions and the students answering singly or in chorus, and (2) the answers alone as composition material, the instructor giving the English and the students then giving the Spanish, orally or in writing. Sometimes the students are paired for a question-and-answer period.
>
> There is something about this use of questions which makes the learning easier than outright memorization. Perhaps it seems natural to the mind to give an answer to a question. Since the questions break the reading sentences into shorter units, the students learn to read material which is rather difficult syntactically, but have less difficult material to put back into Spanish. The better ones

---

[3] The first administration of the aurals at the college, in May, 1946.

learn to string the short sentences into longer ones, while the poorer students still have material which is within their ability.

The first book we use is Keniston's *Learning Spanish*, for three reasons: (1) the constructions are presented in order of frequency, so the students get most practice on the most common ones; (2) the material in the reading section of the lessons is interesting, and the book teaches them to translate well from Spanish to English or, if you prefer, to read and understand Spanish well. Starting at the first of the past school year, we used the question system throughout the book. The second book we used was *A México por automóvil*, and questions were used throughout. The third book was *Noche Oscura en Lima*, with questions used only in the first fifteen lessons; after that we read rapidly to finish the book before the end of the year.

We do considerable chorus work in our classes - reading aloud, answering questions, translating English to Spanish - because in that way each individual gets more practice. An inexperienced instructor can follow the questions closely and give his students a lot of oral work; an experienced instructor will begin with the questions as given and gradually work his students into freer questioning and answering, sometimes on a level that might be called conversation.

The second year classes had had some of the question method in their first year (the last twelve lessons of *Learning Spanish*, twelve lessons of *A México por automóvil*, and fifteen lessons of *Noche Oscura en Lima)*, except for those who entered the course after two years of high-school Spanish. In their second year we put them through House and Mapes' *Shorter Spanish Grammar* to start with. We find that a beginning grammar is better than a review grammar for co-ordinating our own students and those coming from high school. Of course we go through it more rapidly than we would if we were using it with beginning students. After the grammar was well started, we began reading Cano and Saenz' *Easy Spanish Plays*. Enough of the plays were staged in the classroom to give each student a chance to memorize and perform one long part or two or more small parts. Next they read *El Diablo Blanco* and most of Olmstead's edition of Becquer's *Legends, Tales and Poems*. Running concurrently with the readers and with each other were Berkowitz's *Temas de Repaso* and Madrigal and Madrigal's *Invitation to Spanish*. This last book was used in an attempt to maintain what oral proficiency they had previously acquired, with material so simple that it would not take much time either for study or in class. We wanted the oral work without taking much time from the reading. I do not feel that it was very successful.

In 1946-1947, three Spanish groups participated in the testing program: (1) eighty-four first-year students, (2) fifty-eight intermediate (second-year) students, (3) sixteen advanced students.

Group 1. The first-year course met five hours per week for an approximate total of 150 class hours. The students were tested both visually and aurally at the end of the first semester (January), and again at the end of the year (May). The results in mean scores were as follows:

	CO-OPERATIVE					AURAL		
	JAN. (75 HRS.) ADV. P		MAY (150 HRS.) ADV. Q			JAN. (75 HRS.) LOWER B	MAY (150 HRS.) LOWER A	
	MEAN	(%ILE)#	MEAN	(%ILE)#		MEAN	MEAN	(TENTH)$
R*	43.4	(13)	60.9	(77)	I	3.5	4.0	(4)
V	43.7	(11)	59.4	(64)	II	4.7	8.8	(4)
G	45.6	(25)	54.6	(61)	III	7.8	14.0	(6)
T	43.8	(13)	59.2	(68)	T	16.1	26.7	(4)

[*] R = Reading; V = Vocabulary; G = Grammar; T = Total

[#] These percentiles are drawn from *The Co-operative Achievement Tests: A Booklet of Norms*, published by the Co-operative Test Service of the American Council on Education, New York City, May 1938. The present figures are found on pp. 68-71: End-of-year percentile norms in terms of Scaled Scores for College Students who have had no instruction in Spanish in secondary school: Type II (typical liberal-arts) colleges. — In all subsequent reporting of norm percentiles for the Co-operative tests, we refer to the appropriate years-of-study column of the Type II norms as set forth in this booklet.

[$] These tenths are drawn from the decile norms as set forth for the Aural Tests in Chapter IV, pp. 65-70. The figures for the May testing are taken from Spanish Norm-group 1B; figures are not available for the January testing, since Norm-group 1A (less than 90 hours) was not constituted for Form B.

Evidently the reading-objective was not unduly sacrificed for the sake of aural training with this group, who from January to May advanced 15.4 scaled-score points in the total Co-operative score to reach the 68th percentile, and was by then strongest in Reading. (They progressed least in Grammar, on which part they were lowest in May.) While they advanced substantially also on the aural test (especially on the anecdotes), they remained slightly below the median in this field.

Group 2. The second-year group was composed of thirty-five students who had completed one year of Spanish in the college (whom we shall call Subgroup 2.1); and of twenty-three students, all but four of them entering freshmen who had begun their study of the language in secondary school (designated as Subgroup 2.2). The course met four hours per week for a total of 120 hours. The full group was tested at the beginning (October) and at the end (May) of the year. The results were as follows:

	CO-OPERATIVE					AURAL			
	OCT. (150 HRS.*) ADV. P		MAY (270 HRS.) ADV. Q			OCT. (150 HRS.*) LOWER B		MAY (270 HRS.) LOWER A	
	MEAN	(1 YR. %ILE)	MEAN	(2 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	59.4	(71)	66.2	(51)	I	8.4	(6)	8.6	(7)
V	52.6	(39)	63.4	(35)	II	9.9	(6)	13.8	(6)
G	56.0	(65)	66.2	(64)	III	13.6	(5)	18.6	(7)
T	56.7	(60)	66.9	(53)	T	32.0	(6)	41.0	(7)

*For the entering Freshmen, this figure naturally varied with the amount of high-school study.

This group, which began and ended lowest on the Vocabulary section of the Co-operative, was not as high (60th percentile) at the beginning of the second year as the elementary group was at the end of the first (68th percentile). It was the members who began Spanish in the college (Subgroup 2.1), however, who pulled the total mean down; the entering freshmen had a total mean of 57.9, while that of Subgroup 2.1 was 55.6 (this difference, however, did not prove to be significant). By the end of the 270 hours the full group, still relatively weak on Vocabulary, had attained the 53rd percentile of the *two*-year Co-operative norms. At this time Subgroup 2.1 had a Total mean of 64.6, while those who had entered the course from high school scored 70.0, showing a difference of 5.4 points, significant at the 1% level (i.e., with only one chance in a hundred that the difference was due to error in sampling).

Since usually such a significant difference is more likely to appear at the beginning of the year and to level out as all members of the class are subjected to identical instruction, we are led to inquire if other differences between the two subgroups existed to account for this one. A comparison of L-scores on the American Council Psychological Examination (taken at entrance) shows a mean difference in favor of the high-school subgroup of 5.8, which, however, does not prove to be significant. At the same time a comparison of cumulative grade-point averages for the second semester shows a mean difference in favor of the other (college) subgroup of .08, and this, again, is non-significant. In the absence of significant differences in intelligence and in general college achievement between the two subgroups, one person's guess is as good as another's as to why the significant difference appears in the Co-operative results. It can be that the high-school subgroup, whose performance at the beginning of the year was not significantly better, nevertheless did possess a solider grounding and was the better able to profit from the year's instruction. Or it can be that the high-school group, most of whom were freshmen, worked harder in their first year of college Spanish than the others, most of whom were sophomores, worked in their second year. Or it may be that neither of these factors, but another or others which escape our search, account for the difference.

On the first aural test, the subgroup entering from high school also scored higher (total 32.5) than the college subgroup (total 30.2), but the difference was not significant. At the end of the year, the high-school subgroup totaled 42.9 on the aural test as against 39.4 for the college-trained students. Again the difference was non-significant; it is noteworthy, however, that the latter subgroup scored lower despite the fact that twenty-one out of its thirty-five members had already taken the same form of the lower-level aural test one year previously.[4]

These twenty-one compiled a total mean of 45.7 as against 30.9 for the balance of the subgroup. Some of them doubtless recalled certain items (particularly one or another of the anecdotes retold to them by fellow students after the testing), and thus their scores on the 1947 repetition must be recognized as raising the mean score of the subgroup through an extraneous factor; had it not been for this factor, the aural mean for Subgroup 2.1 might have been significantly lower than that of the high-school group, as was the case with the Co-operative test.

Group 3. These were members of a course in Spanish Phonetics and Conversation, described in the college catalogue as "A scientific study of pronunciation and oral expression. Intensive drill in pronunciation with individual use of the phonograph for corrective purposes. The second semester is devoted primarily to conversation and oral reports." Classes were held two hours per week, for a total of 30 hours in the year. Sixteen students were tested in January, and the seven of them who continued through the second semester were tested again in May. The results follow:

---

[4] Their mean scores on that testing: I, 7.2; II, 12.9; III, 13.1: Total, 33.1 (7th tenth, Norm-group 1B).

	CO-OPERATIVE					AURAL			
	JANUARY (300 HRS.) ADV. P		MAY (330 HRS.) ADV. Q			JANUARY (300 HRS.) LOWER B		MAY (330 HRS.) LOWER A	
	MEAN	(3 YR. %ILE)	MEAN	(3 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	74.2	(52)	70.9	(40)	I	13.6	(8)	11.1	(8)
V	71.1	(32)	74.9	(47)	II	15.9	(7)	17.6	(8)
G	72.7	(74)	70.3	(64)	III	20.0	(7)	21.9	(8)
T	75.0	(55)	74.1	(51)	T	49.5	(7)	50.6	(8)

Already in January this group stood at the 55th percentile of the three-year Co-operative norm — which was to be expected of self-selected students taking advanced Spanish because of previous success in, and liking for, language-study. It is curious, however, that the seven finishers were less successful at mid-year than those who discontinued: their Total Co-operative score was only 73.0 (47th percentile). Thereafter, while increasing their Vocabulary scores considerably, they fell off in Reading and Grammar, a situation probably explained by the fact that their second semester was devoted chiefly to oral work and did not continue to exercise those reading skills.

In January the total aural mean of the seven continuing students (42.9) was lower by a decile than that of the full group (49.5) and thus the final mean of 50.6 indicates improvement during the second semester. In noting the decile ranking on these tests, however, one should remember that these students had accumulated many more class hours than most students on whose scores the norms are based.

French

Two French groups were followed through 1946-1947 at College A: (1) fifty-one regular first-year students, and (2) sixteen students taking an accelerated elementary-intermediate course. The two courses are described in the college catalogue as follows:

> (1) Elementary French. — An introduction to spoken French and elementary French reading. Pronunciation, oral work, essentials of

grammar, and the reading of easy materials dealing with the civilization and contemporary affairs of France. Students wishing to develop an ability to understand and speak elementary material spend at least a half hour each day in laboratory work with phonograph records, with locally made transcriptions of lesson materials and in conversational practice with a French-speaking assistant.

(2) Elementary Speaking and Intermediate Reading. — A concentrated course for beginners of approved aptitude, designed to develop a basis for oral facility and a reading ability in two semesters.

Group 1a. This regular elementary group met five hours per week for a total of 75 hours each semester, exclusive of the optional laboratory hours. The testing procedure was the same for the French beginners as for the Spanish, i.e., Co-operative and aural at the end of each semester. The results:

	CO-OPERATIVE					AURAL		
	JANUARY (75 HRS.) ADV. Q		MAY (150 HRS.) ADV. R			JANUARY (75 HRS.) LOWER B	MAY (150 HRS.) LOWER A	
	MEAN	(%ILE)	MEAN	(%ILE)		MEAN	MEAN	(TENTH)
R	45.5	(19)	59.1	(66)	I	4.2	5.1	(6)
V	57.2	(39)	60.0	(52)	II	10.7	13.1	(7)
G	52.9	(45)	60.8	(76)	III	6.1	10.2	(5)
T	51.9	(32)	61.0	(67)	T	21.0	28.4	(7)

From January to May the group advanced 9.1 points in the Total Co-operative score to reach the 67th percentile, and was strongest in Grammar. By comparison with Spanish Group 1, these French beginners were more advanced in the reading skills at mid-year but were at the same level at the end of the year. In aural ability, however, the group placed in the 7th tenth at the May testing as compared with the Spanish group in the 4th. Differences in intelligence and general college achievement between the French and Spanish beginning groups favor the French[5]; but it is doubtful if these differences account for the higher aural performance so long as the corresponding Co-operative scores are no higher, since both tests yield about the same correlation with intelligence.

---

[5] The ACE L-scores average 74.6 and 67.9; the grade-point averages are 1.65 and 1.37, respectively.

This elementary French group also offers the opportunity to study the effect of previous experience with the language. Thirty-eight who had had no previous experience with French, we will call Subgroup la.1. The other thirteen members, although enrolled in the elementary course, had already begun to study French in high school.[6] These we will call Subgroup la.2. In examining differences between the Co-operative and aural scores of these subgroups, we find that at midyear the students with high-school experience made a Co-operative Total mean of 57.5 as against 49.9 for the raw beginners; on the aural they scored 26.2 as against 19.1. Both these differences are significant at the 1% level. It is quite possibly the fairly liberal sprinkling of students with previous experience in this group which accounts for the higher reading achievement at 75 hours in French than in Spanish, for in Spanish only six of the eighty-four beginners had had high-school experience.

At year's end the subgroup with previous experience (la.2) was still ahead on the Co-operative by 64.1 to 59.9 and on the aural test by 29.2 to 28.4 but the differences are no longer significant. This is the usual situation in which elementary students having previous experience with the language fail to maintain their advantage over the raw beginners throughout a year. In practical terms this also means that raw beginners are scarcely justified in complaining that they are handicapped in competition with the partially initiated members of their group.

Group 1b. Testing of this accelerated elementary-intermediate group was the same as for Group la with the following results:

---

[6] One had three years, four had two years, and eight had one year. Their average grade for these years was B, which does not suggest that they began the language again in college because of earlier failure.

	CO-OPERATIVE					AURAL			
	JANUARY (90 HRS.) ADV. Q		MAY (180 HRS.) ADV. R			JANUARY (90 HRS.) LOWER B		(MAY (180 HRS.) LOWER A	
	MEAN	(1 YR. %ILE)	MEAN	(2 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	42.6	(13)	65.4	(42)	I	2.9	(3)	3.3	(4)
V	49.8	(16)	60.3	(10)	II	8.1	(2)	9.1	(4)
G	52.0	(40)	59.8	(44)	III	6.3	(4)	10.0	(5)
T	47.9	(19)	63.1	(28)	T	17.3	(2)	22.4	(4)

Though this course aimed at the development of an "intermediate reading" ability in two semesters, the Co-operative scores do not indicate that this objective was successfully attained. In fact it is doubtful that it could be reached in a total of only 30 more hours of class instruction than the elementary group received. At midyear the group stood only at the 19th percentile of the one-year norms in the Co-operative, and reached that level chiefly on the strength of their showing in Grammar. At the end of the year their reading skills had improved 15.2 points to reach the 28th percentile of the two-year norms. Incidentally, although they managed to surpass the regular elementary group at the end of the year, at midyear they scored lower on all parts of the Co-operative test, even though they had had 15 more hours of instruction. This might conceivably be due to a proportionally greater emphasis on oral-aural work during the first semester, with a consequent sacrifice of reading development in early stages. However, this difference of emphasis failed to produce compensatingly higher aural scores; on the contrary, the group scored lower on the aural tests than the regular beginners both at midyear and at the end, when they placed only in the 4th tenth, as contrasted with the 7th for the beginners who had 30 less hours. This failure of the more intensive French Group 1b as contrasted with Group 1a cannot be explained on the basis of such factors as intelligence or application, for the differences between the means on the ACE L-score and on cumulative grade-point average are not significant.

Russian

Two small Russian groups -- the only ones receiving instruction in this language at College A — were also tested: (1) twelve elementary students, and (2) a dwindling handful of second-year students. The college catalogue announces the respective courses as follows:

> 1. An introduction to the Russian language with such training in grammar as is required for reading a highly inflected language. The main purpose of the course is to provide the basis for a reading knowledge of current writing (newspapers and non-technical magazines) and of current literature, but there will also be practice in using the idiom of everyday speech.
>
> 2. A continuation of the study of the Russian language and its literature.

Group 1. This elementary group was tested only at the end of the year's course, after 150 hours at five per week. The reading and aural scores:

	INVESTIGATION'S TEST			AURAL LOWER A	
	MEAN*	(TENTH)		MEAN	(TENTH)
R	13.7	(3)	I	4.7	(3)
V	5.3	(1)	II	10.8	(4)
G	9.8	(3)	III	16.8	(5)
T	28.8	(2)	T	32.3	(5)

*Scores for the Investigation's "Co-operative-type' Russian test are raw, not scaled. As with the aural tests, place in the distribution is marked by deciles. The decile table for this test is found in Appendix A.

Scores on the visual test are low for one year of study, but the aural scores are median.

Group 2. Six students began a second year of Russian, after diverse amounts of elementary instruction in 1945-1946. Only three of them had taken the full 150-hour beginning course, completing twenty-two lessons in the Semeonoff grammar and making concurrent use of two elementary readers. The other three had met with the instructor only one evening a week (not for credit) and had covered much less ground; one of these, having begun Russian in Australia some years before, had not resumed his study at College A until midyear. In 1946-1947 these six students began as a single "second-year" group, meeting four hours per week. Their work was divided into three parts: grammar review, conversation, and reading, about half the time being devoted to the latter. The conversation work was handled by a native Russian.

The six were tested in January. For this small group only the total means are reported here:

	INVESTIGATION'S TEST			AURAL LOWER A	
	MEAN	(TENTH)		MEAN	(TENTH)
T	31.7	(3)	T	16.7	(1)

While the score for the reading skills is higher than that of the elementary group, the aural score is far lower. Two facts account for this discrepancy: (1) these students received a very small quantity of ear training in their introductory course, and (2) none of them was of Slavic family background, as against four such students in Group 1.

Three of the six continued the Russian course to the end of the year and were tested again in May with the same forms of the two tests. Their total mean scores were: R-V-G test, 39.0 (4th tenth); aural, 18.0 (2nd tenth). Their progress during the second semester is thus observed to have been relatively slight.

Further Studies

The mass of supplementary data furnished us by College A makes possible, through the use of correlation techniques, some further studies which illumine certain aspects of language teaching, not only at College A but in general. The complete set of coefficients appears as Table IX and is based on Spanish Groups 1 and 2 and French Group 1a.

TABLE IX

College A: Correlations

		CO-OPERATIVE Q	ACE L-SCORE	SECOND SEMESTER GRADE	GRADE-POINT AVERAGE	CO-OPERATIVE P	AURAL LOWER B
Spanish Group 1 (N=84)	ACE L-SCORE*			.44	.36		
	CO-OPERATIVE Q		.54	.73		.72	
	AURAL LOWER A	.49	.42	.22			.60

Partial Correlations:

1. Co-operative Q and Aural Lower A, with ACE L-score held constant: .34
2. Co-operative Q and Second Semester Grade, with ACE L-score held constant: .65
3. Aural Lower A and Second Semester Grade, with ACE L-score held constant: .04
4. Aural Lower A and Second Semester Grade, with Co-op. Q held constant: -.23

		CO-OPERATIVE Q	ACE L-SCORE	SECOND SEMESTER GRADE	GRADE-POINT AVERAGE	CO-OPERATIVE P	AURAL LOWER B
Spanish Group 2 (N=56)	ACE L-SCORE			.26	.31		
	CO-OPERATIVE Q		.36	.75		.69	
	AURAL LOWER A	.68	.19	.40			.77

Partial Correlations:

1. Co-operative Q and Aural Lower A, with ACE L-score held constant: .67
2. Co-operative Q and Second Semester Grade, with ACE L-score held constant: .73
3. Aural Lower A and Second Semester Grade, with ACE L-score held constant: .37
4. Aural Lower A and Second Semester Grade, with Co-op. Q held constant: -.23

		CO-OPERATIVE R	ACE L-SCORE	SECOND SEMESTER GRADE	GRADE-POINT AVERAGE	CO-OPERATIVE Q	AURAL LOWER B
French Group 1a (N-51)	ACE L-SCORE			.41	.43		
	CO-OPERATIVE R		.43	.70		.58	
	AURAL LOWER A	.44	.40	.51			.69

Partial Correlations:

1. Co-operative R and Aural Lower A, with ACE L-score held constant: .32
2. Co-operative R and Second Semester Grade, with ACE L-score held constant: .64
3. Aural Lower A and Second Semester Grade, with ACE L-score held constant: .41
4. Aural Lower A and Second Semester Grade, with Co-op. R held constant: .31

*Language Score on the American Council on Education's Psychological Examination.

The first issue is the relation between the reading and the aural skills. At midyear the correlations between the aural tests (B forms) and the Co-operatives (Spanish P and French Q) range from about .60 to .70. These coefficients are slightly higher than usual, but it scarcely seems worth while to seek an explanation since at the end of the year the correlations (.44-.68) between the A-forms of the aurals and Spanish Q and French R fall within the usual range.

An obvious further question in this regard is how much is basic intelligence the common factor which tends to determine the scores on both tests. Taking the L-score of the ACE Psychological Examination as the measure of this trait and holding its influence constant by partial correlation, we find that the coefficients are lowered by .01-.15, but, in the main, such relation as exists between the two sets of abilities does not depend primarily on their common relation to intelligence.

The relative emphasis given the different skills at College A can be clearly seen by examining the correlations between the grades assigned at the end of the second semester and the students' scores on the two tests at that same time. The coefficients between the final grade and the Co-operative score are the highest in the entire table, ranging from .70-.75. Even when intelligence (as measured by the ACE L-score) is partialed out, this relation remains high (.64-.73). In short, College A, like many other schools which stress reading ability, grades students on their ability to read, write, and translate, and hence on the skills measured by the Co-operative tests. On the other hand, aural ability figures much less in the final grade, with the coefficients ranging from .22 to .51. If intelligence is partialed out, these figures drop to the range .04-.41. Perhaps the most revealing of all is the further fact that, if the abilities measured by the Co-operative test are partialed out, then the correlations between the aural score and the final grade fall still lower, two of them being -.23.

Another point illuminated by these figures is the consistency with which students develop the various skills. This evidence lies in the correlations between the two administrations of each type of test. For both tests these coefficients are of about the order .60-.70. This is a fairly high relation; but the familiar cases of the student who "doesn't find himself until late in the year" or who "tries to live all year off his hard work the first few weeks while his good resolutions lasted" keep the progress of the class as a whole from being uniform in gaining either set of skills.

Finally, it is worth noting that the measure of verbal intelligence is about as good a predictor of elementary language grades as it is of other school achievement (as indicated by the grade-point average) and that the degree of relation (roughly .30-.40) is typical.

Since College A is the first institution we have studied in detail and since it presents a typical picture, we have commented on these points at some length. Hereafter we shall mention only those instances which deviate markedly.

Summary of College A

College A is the type of school often considered fairly representative of small liberal arts colleges. This typicalness is substantiated by the descriptions and data we have just presented. The instruction is "conventional" in the sense in which we use the term. The general procedure is that followed in many colleges, with, as usual, such modifications as are suggested by local conditions or the experience of the staff. Attainment in the reading skills is average or better, and some superiority is to be expected at an institution where the staff is sufficiently interested to offer its work for examination. These reading skills receive chief emphasis in the course, as is indicated both by the formal statements and by the heavy weighting of these abilities in awarding the final grades. The level of aural comprehension also runs slightly higher than average and constitutes the most unusual feature of the picture. A fairly large number of class hours, slightly more than usual emphasis on the aural-oral skills, and somewhat less vocabulary trouble for the students vis-a-vis the aural tests (as compared with students using some of the aural-oral texts) are factors which may account for this performance. The relations of the skills to each other, to intelligence, and to general college achievement are all fairly typical.

College B

We arranged a program of collaboration with College B (a medium-sized city university) chiefly because its department of Romance languages had been operating, since the fall of 1944, a series of experimental intensive courses at different levels in addition to retaining all its conventional three-hour-per-week courses. We tested both the experimental and the regular classes in 1945-1946, and obtained a large body of supplementary data concerning many of the students' intelligence, high-school language study, success in college language courses, and the like. Interesting comparisons were expected to issue from this data, but all the experimental groups proved to be so small that no significant comparisons could be made between them and the regular groups. Furthermore, local considerations imposed the testing of both types at the same calendar time, instead of at identical amounts of accumulated contact hours, the only point at which comparisons could be meaningful. Consequently we shall discuss the two programs separately, the conventional groups here, and the experimental ones in the following chapter.

Our testing program covered several groups of conventional students at the intermediate and advanced levels in French and at the intermediate level in Spanish.

French

Group 2a. One French group was tested toward the end of its intermediate course. Many of the students had entered the course from high school; the balance, who had begun the study of French in the university, had accumulated 90 hours in the elementary course and were adding 90 more in the intermediate for a total of 180 hours. On that scale the group was tested at 145 hours, in October 1945, with the following results:

Group 2a (N=50)

	CO-OPERATIVE			AURAL	
	OCT. 1945 (145 HRS.) ADV. R			OCT. 1945 (145 HRS.) UPPER A	
	MEAN	(2 YR. %ILE)		MEAN	(TENTH)
R	64.7	(40)	I	6.4	(4)
V	66.9	(28)	II	16.4	(5)
G	63.9	(61)	III	3.1	(6)
T	66.6	(42)	T	25.9	(4)

Groups 2b and 2c. Two other French groups were tested twice — once shortly after the beginning, and again toward the end, of their intermediate course. As is the case of Group 2a, some of these students had entered the course directly from high school while others had formerly completed the elementary college course. The results of the two testings, made at slightly different points on the 180-hour scale, and using different levels of the aural test in the respective groups, were as follows:

Group 2b (N=76)

	CO-OPERATIVE					AURAL			
	OCT.1945(105 HRS.) ADV. R		APR.1946(160 HRS.) ADV. R			OCT.1945(105 HRS.) UPPER A		APR.1946(160 HRS.) UPPER B	
	MEAN	(1 YR. %ILE)	MEAN	(2 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	62.8	(78)	71.8	(67)	I	5.6	(4)	4.0	(4)
V	66.6	(77)	71.6	(47)	II	15.8	(5)	11.2	(5)
G	66.0	(89)	68.0	(76)	III	2.3	(5)	4.0	(5)
T	66.7	(85)	72.7	(65)	T	23.7	(4)	19.2	(5)

Group 2c (N=33)

	FEB. 1946 (95 HRS.) ADV. R		JUNE 1946 (140 HRS.) ADV. R			FEB. 1946 (95 HRS.) LOWER B		JUNE 1946 (140 HRS.) LOWER A	
	MEAN	(1 YR. %ILE)	MEAN	(2 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	58.3	(63)	68.9	(61)	I	8.0	(7)	8.4	(7)
V	66.3	(76)	71.0	(45)	II	14.7	(7)	14.8	(6)
G	62.7	(83)	66.9	(72)	III	7.6	(5)	12.7	(6)
T	63.8	(76)	71.0	(59)	T	30.3	(7)	35.9	(6)

Both Groups 2b and 2c stood above the two-year Co-operative median at the second testing in all scores for the Co-operative except Vocabulary, and all groups were strongest in Grammar. On the aural test the sub-median total score of Group 2b at the first testing may be due in part to the fact that the upper-level test was too difficult for students with only 105 hours.

Group 3a. This was a small group of seventeen students, all of whom entered the intermediate course with three or four years' study of high-school French in which they had made superior records. They took essentially the same course as the intermediate groups reported above, but were taught in a separate section in which more reading was done and more ground covered in all phases of the course. They thus qualify, in a sense, as an advanced group; consequently their scores on the tests, considerably higher than those of the preceding groups, are measured at the end in terms of the third-year norms. In reporting the points at which testing was done, we use a scale of HS (High School) + 90 hours. The scores:

Group 3a (N=17)

	CO-OPERATIVE					AURAL			
	OCT. 1945 (HS+15 HRS.) ADV. R		APR. 1946 (HS+70 HRS.) ADV. R			OCT. 1945 (HS+15 HRS.) UPPER A		APR. 1946 (HS+70 HRS.) UPPER B	
	MEAN	(2 YR. %ILE)	MEAN	(3 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	79.9	(89)	85.6	(85)	I	13.9	(5)	8.6	(5)
V	76.8	(68)	79.2	(40)	II	23.1	(5)	18.4	(6)
G	67.8	(75)	70.8	(73)	III	2.7	(3)	7.3	(5)
T	76.9	(80)	82.1	(72)	T	39.7	(5)	34.3	(6)

Like other intermediate groups, these students ended weakest in Vocabulary, but unlike them they finished strongest in Reading rather than in Grammar.

Group 3b. This was a small group of eleven students taking a survey of literature. Tested at what is roughly calculated on our scale as 235 hours, they produced the following scores:

Group 3b (N=11)

	CO-OPERATIVE			AURAL	
	APR. 1946 (235 HRS.) ADV. R			APR. 1946 (235 HRS.) UPPER B	
	MEAN	(3 YR. %ILE)		MEAN	(TENTH)
R	81.9	(75)	I	7.7	(4)
V	79.4	(40)	II	17.1	(5)
G	74.5	(84)	III	4.3	(3)
T	81.5	(70)	T	29.1	(4)

The low aural scores of this group may possibly be attributable to their relatively slight use of the spoken language in a literature course.

Effect of previous study of French for Groups 2b and 3a. The scores of Groups 2b and 3a on the Co-operative and Aural scores of the October testing were analyzed to determine the effect of varying amounts and kinds of previous study of French. There were three categories of previous experience: (1) no high-school study with the elementary course taken at the University; (2) two years of high-school French; (3) three or four years of high-school French. Means of these three subgroups on the Co-operative and aural *totals* were as follows:

	CO-OPERATIVE	AURAL
1. One year of college French (9 cases from Group 2b)	59.6	15.8
2. Two years of high-school French (35 cases from Group 2b)	65.4	24.4
3. Three or four years of high-school French (39 cases from Group 2b, 16 from 3a)	72.3	30.0

All the differences between the means of Categories 1-2, 1-3, or 2-3 are significant except that between 1 and 2 on the Co-operative total. In other words, two years of high-school French are about the equal of elementary French at this university as far as the reading skills are concerned but for aural comprehension give a better base than does the university work. Likewise these data indicate that three or four years of high-school experience produce better results than only two, but whether the difference justifies the time is another question.

Spanish

Group 2a. One Spanish group was tested toward the end of its intermediate course. Many of the students had entered the course from high school; the balance, who had begun the study of Spanish in the university, had accumulated 90 hours in the elementary course and were adding 90 more in the intermediate for a total of 180 hours. On that scale the group was tested at 145 hours, in October 1945, with the following results:[7]

	CO-OPERATIVE			AURAL	
	Group 2a (N=63)				
	OCT. 1945 (145 HRS.) ADV. Q			OCT. 1945 (145 HRS.) UPPER A	
	MEAN	(2 YR. %ILE)		MEAN	(TENTH)
R	65.8	(50)	I	9.0	(5)
V	65.0	(42)	II	22.1	(6)
G	67.0	(68)	III	3.4	(4)
T	67.5	(55)	T	34.5	(5)

Groups 2b and 2c. Two other Spanish groups were tested twice — once shortly after the beginning, and again toward the end, of their intermediate course. Approximately one half of these students had entered the course directly from high school, while the other half had completed the elementary college course. The results of the two testings, made at slightly different points on the 180-hour scale, and using different levels of the Aural Test in the respective groups, were as follows:[8]

---

[7] Cf. the identical testing situation for French Group 2a.

CO-OPERATIVE AURAL

Group 2b (N=127)

	OCT.1945(105 HRS.) ADV. Q		APR.1946(160 HRS.) ADV. Q			OCT.1945(105 HRS.) UPPER A		APR.1946(160 HRS.) UPPER A	
	MEAN	(1 YR. %ILE)	MEAN	(2 YR. %ILE		MEAN	(TENTH)	MEAN	(TENTH)
R	59.9	(74)	66.8	(54)	I	5.5	(4)	8.1	(5)
V	59.3	(65)	66.3	(48)	II	18.8	(4)	24.9	(6)
G	61.9	(85)	66.8	(67)	III	3.7	(5)	4.8	(6)
T	61.5	(77)	68.4	(59)	T	28.0	(4)	37.8	(6)

Group 2c (N=16)

	FEB.1946(95 HRS.) ADV. Q		JUNE 1946(140 HRS.) ADV. Q			FEB.1946(95 HRS.) LOWER B		JUNE 1946(140 HRS.) LOWER A	
	MEAN	(1 YR. %ILE)	MEAN	(2 YR. %ILE		MEAN	(TENTH)	MEAN	(TENTH)
R	60.0	(74)	66.6	(54)	I	8.6	(7)	9.9	(7)
V	58.2	(61)	67.0	(50)	II	11.0	(6)	14.4	(6)
G	59.5	(77)	63.3	(53)	III	18.7	(7)	20.4	(7)
T	60.2	(73)	67.3	(54)	T	38.3	(7)	44.7	(8)

Just as in the case of French Groups 2b and 2c, both Spanish groups stood above the two-year Co-operative median at the second testing in all parts except Vocabulary, and Groups 2a and 2b (though not 2c) were strongest in Grammar.

Again as in French, Group 2b's sub-median total on the aural at the first testing may result partially from taking the upper-level test at only 105 hours. The same form of the aural test was repeated at the second testing, which can account for the sharper rise in Spanish than in French, where alternate forms were used.

Effects of previous study of Spanish in Group 2b. Group 2b was analyzed to determine the effect of varying amounts and kinds of previous study of Spanish on the Co-operative and aural scores of the October testing. There were three categories of previous experience: (1) no high-school study — elementary course taken at the University; (2) two years of high-school Spanish; (3) three years of high-school Spanish. Means if these three subgroups on the Co-operative and aural *totals* were as follows:

---

[8] Cf. the identical testing situation for French Groups 2b and 2c.

	CO-OPERATIVE	AURAL
1. One year of college Spanish (57 cases)	58.6	27.0
2. Two years of high-school Spanish (38 cases)	62.8	27.5
3. Three years of high-school Spanish (24 cases)	66.5	30.6

Here the situation is different from that in French. Significant at the 1% level are the differences between the means of Categories 1-2 and 1-3 on the Co-operative total; significant at the 5% level is that between Categories 2-3 on the Co-operative total. On the Aural total, none of the differences between means proved to be significant. In making comparisons with French, one should note that none of the Spanish students in Category 3 had had four years. Thus one extra year of high-school work makes some difference, though not a very significant one.

Supplementary Data and Correlations

For certain of the groups covered by the testing program at College B, we secured two bodies of supplementary data: (1) the individual students' decile scores on the Ohio State University Psychological Examination, taken at entrance; (2) final grades assigned in that semester of the language course during which our tests were administered. The following are the means, by groups:

	OHIO PSYCH. (LOCAL DECILE)	FINAL GRADE
French Group 2a	5.5	B -
French Group 2b	5.9	C +
French Group 3a	7.7	B
Spanish Group 2a	5.9	B -
Spanish Group 2b	5.9	C +

From these figures it is clear that all groups are about the same in intellectual ability (and slightly higher than the local average) except for French 3a, which is well above average and was given a better grade. For these five groups the coefficients of various correlations of the test scores with each other and with the two types of supplementary data are shown in Table X.[9] The picture is so similar to that already seen at College A that only two points merit special comment. Here the correlations between the aural and Co-operative scores are slightly lower than at College A, but they still fall within what appears to be the usual range of these coefficients, roughly .25-.70. The other point is that French 3a, the group with extended high school experience, is typical of most advanced courses in that grades have little relation to aural skill and the correlation between aural score and final grade is very low.

College C

At College C, a medium-sized privately endowed college, we were unable to arrange for a testing program which would follow the progress of certain homogeneous groups. Nevertheless, the aural tests were used extensively at this college in both 1945-1946 and 1946-1947, and the Co-operative tests

---

[9] The number of cases in each group appears slightly reduced here, as a few students had not taken the Ohio Psychological Tests and thus had to be omitted from the correlations.

TABLE X

College B: Correlations

		CO-OPERATIVE R	OHIO PSYCH.	SEMESTER GRADE
French Group 2a (N=43)	OHIO PSYCH.			.30
	CO-OPERATIVE R		.32	.40
	AURAL UPPER A	.46	.35	.30

Partial Correlations:

1. **Co-operative R and Aural Upper A, with Ohio score held constant: .39**
2. **Co-operative R and Semester Grade, with Ohio score held constant: .34**
3. **Aural Upper A and Semester Grade, with Ohio score held constant: .22**
4. **Aural Upper A and Semester Grade, with Co-op. R held constant: .14**

		CO-OPERATIVE R	OHIO PSYCH.	SEMESTER GRADE
French Group 2b (N=69)	OHIO PSYCH.			.34
	CO-OPERATIVE R		.44	.66
	AURAL UPPER A	.46	.37	.35

Partial Correlations:

1. **Co-operative R and Aural Upper A, with Ohio score held constant: .36**
2. **Co-operative R and Semester Grade, with Ohio score held constant: .60**
3. **Aural Upper A and Semester Grade, with Ohio score held constant: .32**
4. **Aural Upper A and Semester Grade, with Co-op. R held constant: .07**

		CO-OPERATIVE R	OHIO PSYCH.	SEMESTER GRADE
French Group 3a (N=16)	OHIO PSYCH.			.16
	CO-OPERATIVE R		.41	.52
	AURAL UPPER A	.36	.72	.05

Partial Correlations:

1. **Co-operative R and Aural Upper A, with Ohio score held constant: .10**
2. **Co-operative R and Semester Grade, with Ohio score held constant: .50**
3. **Aural Upper A and Semester Grade, with Ohio score held constant: -.10**
4. **Aural Upper A and Semester Grade, with Co-op. R held constant: -.17**

TABLE X

(continued)

		CO-OPERATIVE Q	OHIO PSYCH.	SEMESTER GRADE
Spanish Group 2a (N=58)	OHIO PSYCH.			.07
	CO-OPERATIVE Q		.10	.38
	AURAL UPPER A	.29	.23	.23

Partial Correlations:

1. Co-operative Q and Aural Upper A, with Ohio score held constant: .28
2. Co-operative Q and Semester Grade, with Ohio score held constant: .38
3. Aural Upper A and Semester Grade, with Ohio score held constant: .23
4. Aural Upper A and Semester Grade, with Co-op. Q held constant: .14

		CO-OPERATIVE Q (OCTOBER)	OHIO PSYCH.	SEMESTER GRADE
Spanish Group 2b (N=118)	OHIO PSYCH.			.26
	CO-OP. Q (OCT.)		.42	.58
	AURAL UP. A (OCT.)	.43	.35	.26

Partial Correlations:

1. Co-operative Q and Aural Upper A, with Ohio score held constant: .33
2. Co-operative Q and Semester Grade, with Ohio score held constant: .27
3. Aural Upper A and Semester Grade, with Ohio score held constant: .19
4. Aural Upper A and Semester Grade, with Co-op. Q held constant: .17

were administered at the end of each year, and the results were submitted to us for comparison with the aural achievement. During the first of the two years no experimental courses were in operation at the college; all courses were of the conventional eclectic type, and we present the test results in this chapter as representative of such courses in an institution which enjoys excellent national reputation. During the second year (1946-1947), experimental courses were inaugurated in French and Spanish. These, as well as the parallel conventional courses, were covered by our testing, and the results for that year appear in Chapter VI.

In 1945-1946 Form A of the aural tests was given to all classes having sufficient hours at the end of the first semester, and Form B was administered at the end of the year. Thus numbers of individual students were tested aurally twice; but because of constant shifting of enrollment from semester to semester — some students discontinuing, others entering from high school, still others (veterans, for example) resuming the language after a lapse — no single group maintained a constant personnel through even two semesters. We tabulate below the scores for the June testing only; the full table in Appendix B indicates, in a series of notes to each group, the scores made by certain members on either a previous or a subsequent aural testing.

All groups in both French and Spanish met in class three hours per week for a total of approximately 45 hours per semester. In either language, Group 1 had completed one year, Group 2 had completed one and one-half years, Group 3 had completed two years, and Group 4 (made up of students taking various advanced courses) had completed at least two and one-half years. The results by groups:

		CO-OPERATIVE			AURAL	
		ADVANCED FORMS*			LOWER B	
		MEAN	(%ILE)		MEAN	(TENTH)
Spanish	R	60.1	(74)	I	5.7	(5)
Group 1§	V	59.1	(64)	II	8.5	(5)
90 hours	G	66.7	(94)	III	13.6	(5)
(N=56)	T	63.4	(82)	T	27.8	(5)
French	R	63.9	(38)	I	6.9	(6)
Group 2	V	69.7	(38)	II	13.8	(5)
135 hours	G	61.2	(39)	III	10.5	(6)
(N=53)	T	66.5	(41)	T	31.1	(6)

*Form P in Spanish, Form Q in French.

§For reasons not ascertained by us, the second-semester French group was not given the aural test.

		CO-OPERATIVE			AURAL	
		ADVANCED FORMS			LOWER B	
		MEAN	(%ILE)		MEAN	(TENTH)
Spanish	R	66.5	(53)	I	9.8	(5)
Group 2	V	64.3	(40)	II	12.8	(5)
135 hours	G	69.2	(76)	III	17.6	(6)
(N=35)	T	68.6	(59)	T	40.2	(6)
French	R	65.5	(43)	I	4.1	(4)
Group 3	V	70.1	(40)	II	9.5	(5)
180 hours	G	62.0	(52)	III	3.3	(4)
(N=23)	T	67.0	(43)	T	16.9	(4)
Spanish	R	69.4	(65)	I	11.1	(5)
Group 3	V	67.9	(54)	II	7.9	(6)
180 hours	G	69.7	(78)	III	5.1	(4)
(N=25)	T	70.8	(68)	T	24.1	(6)
French	R	82.1	(75)	I	9.6	(6)
Group 4	V	79.6	(41)	II	17.1	(5)
225 hrs. or more	G	71.1	(74)	III	6..7	(5)
(N=27)	T	80.6	(66)	T	33.3	(6)
Spanish	R	73.6	(50)	I	15.1	(5)
Group 4	V	77.4	(56)	II	16.9	(6)
225 hrs. or more	G	75.3	(81)	III	8.8	(5)
(N=30)	T	78.0	(67)	T	40.9	(6)

For all three groups where both languages are represented - (namely 2, 3, and 4), the Spanish norm percentiles are higher than the French on the Co-operative Test. Spanish Group 2, for example, had already at midyear surpassed the median of the two-year norms, while French Group 2 had not. The only instance of a higher percentile in French is that of the Reading in Group 4. An outstanding feature of the Spanish picture (similar to that in Spanish at College B) is the predominantly high level of performance in Grammar, apparent in all groups, which gives direct evidence of precise training in that aspect of the language. Reading development did not suffer, however, all scores being above the median in that department, though the Vocabulary scores tend to run low.

On the other hand, as is neatly demonstrated in a series of Co-operative scores from College G (see below, p. 152), a high degree of reading success can also be achieved with a minimum of grammatical analysis and a maximum of vocabulary control.

The various aural scores for College C offer no noteworthy interrelations, and none of them deviate more than one decile from the median of the appropriate norm-group. It must be admitted that one of the reasons for this uniformity is the fact that College C was by far the largest single contributor to the norm-distributions for the B forms of the aural tests.

Correlations between Visual and Aural Tests

The following are the coefficients of correlation between the Co-operative and the aural tests (total scores) at College C:

Spanish - Group 1:	**.24**
French - Group 2:	**.63**
Spanish - Group 2:	**.65**
French - Groups 3-4 (combined)	**.58**
Spanish - Groups 3-4 (combined)	**.51**

The Group 1 correlation is characteristically low; these students being at a stage where their aural comprehension was relatively undeveloped. For Group 2, with only one-half year more of instruction, the rise above Group 1 is surprisingly sharp, and the figures are unusually close together in the two languages. The slight drop in Groups 3-4 is only apparent and not real.

College D

At College D, a small liberal-arts college, although oral-aural work was regarded as being of secondary importance as compared with the visual skills, a plan was developed for

1945-1946 whereby a maximum of opportunity for aural practice could be provided through laboratory work with the phonograph. Elementary courses in the three languages offered at this college — French, German, and Spanish — held four class meetings per week; in addition, the students attended two supervised laboratory hours each week for work with transcriptions of locally recorded materials related to their course of study. Intermediate (second-year) classes operated under this same plan during the first semester, but at midyear discontinued the two hours of laboratory.

The elementary groups in all three languages were tested at the end of the first semester, when they had accumulated 90 contact hours (inclusive of the laboratory hours). The results:

	CO-OPERATIVE			AURAL	
	ADV.*			LOWER A	
	MEAN	(%ILE)		MEAN	(TENTH)
French (N=77)					
R	48.0	(26)	I	3.2	(4)
V	60.0	(52)	II	8.5	(3)
G	59.8	(73)	III	9.2	(5)
T	56.3	(47)	T	20.9	(4)
German (N=30)					
R	51.9	(34)	I	4.8	(5)
V	54.9	(44)	II	6.1	(4)
G	52.9	(51)	III	6.5	(4)
T	53.2	(40)	T	17.4	(5)
Spanish (N=20)					
R	52.3	(42)	I	3.2	(4)
V	51.2	(32)	II	8.3	(3)
G	54.3	(57)	III	10.4	(3)
T	52.7	(44)	T	22.0	(3)

*Form Q in French, P in German, and Q in Spanish.

In all three languages the highest percentiles on the Co-operative tests are in Grammar. At midyear the totals are all reasonably near the median for one year of study. These same groups were re-tested at the end of the year; but through an oversight the Co-operative results were not made available to us, and the aural scores are of minor interest because not only was the same form of the test (Lower Level A) repeated, but also the groups had become substantially reduced in size. The scores from this testing appear in Appendix B; it is sufficient here to record that in the total, the French group progressed to the 5th tenth, the German to the 8th tenth, and the Spanish to the 7th tenth, of Norm-group 1B.

Correlation coefficients between total scores on the Co-operative and the Aural Tests for the midyear testing were: for French .67; for German .80; for Spanish .40. The unusually high coefficient in German is explained by the cases of two students who made inexplicably high scores on both tests; with these two cases excluded, the figure is .66. After only one semester's instruction, even at the rate of six hours per week, the French and German correlations are surprisingly high.

The intermediate groups in the three languages were also tested, but they took the lower-level aural examination only at the end of the year when, as in the case of the elementary groups, their Co-operative scores were not available to us. In French and Spanish these groups placed in the same tenth of their norm-group as did the elementary groups on their end-of-year re-take of the Aural, while in German they placed only in the 5th tenth as contrasted with the beginners who had jumped to the 8th. Full scores of the intermediate groups appear in Appendix B.

College E

Basic language courses at College E, a large city university, had been modified by current influences to the extent of supplementing the traditional three hours of class per week by a single drill-session of one hour and a half. This drill-period was devoted to oral and aural work, mainly "conversation", under the direction of a native speaker. It was first instituted with the elementary classes in 1944-1945, and was put into the intermediate course the following year. Thus the elementary groups, when tested in May 1946, had accumulated approximately 90 class-hours plus 45 drill-hours for a total of 135 running hours; the intermediate groups had accumulated 90 hours (without benefit of drill-sessions) the first year, and had then added 135 hours under the new arrangement for a total of 225. The results of the testing program, by French and Spanish elementary and intermediate groups:

CO-OPERATIVE			AURAL		
French Group 1 135 hours (N=107)					
	ADV.*			LOWER A	
	MEAN	(%ILE)		MEAN	(TENTH)
R	57.1	(59)	I	3.8	(5)
V	56.5	(35)	II	10.2	(4)
G	56.0	(57)	III	7.9	(4)
T	57.7	(55)	T	21.9	(4)
Spanish Group 1 135 hours (N=129)					
R	61.8	(80)	I	5.9	(6)
V	57.5	(60)	II	11.3	(6)
G	54.2	(57)	III	13.7	(6)
T	58.7	(68)	T	30.9	(5)
French Group 2 225 hours (N=47)					
				UPPER A	
R	80.0	(89)	I	9.5	(6)
V	77.2	(69)	II	17.1	(5)
G	71.3	(86)	III	4.0	(7)
T	78.9	(86)	T	30.5	(6)
Spanish Group 2 225 hours (N=74)					
R	70.5	(67)	I	9.9	(6)
V	75.8	(82)	II	21.9	(6)
G	69.5	(72)	III	4.9	(6)
T	74.2	(79)	T	36.7	(6)

*Form R for French, Form Q for Spanish

These results again reflect the typical situation, characterized by scores well above the median in the reading skills, but by only average, or slightly better, aural achievement. The intermediate groups, with the two languages very near one another in general level of achievement performed better than the elementary groups both visually and aurally. The comparative lack of success for French Group 1 in both areas may be due to various incidental factors (e.g., intelligence); no data is available on which to base any explanation.

College F

Our program at College F, a large city college, was limited to 128 students of elementary German. They are of interest chiefly as a conventionally taught group who were tested at the end of their first semester after only about 50 hours of class instruction. The scores:

	CO-OPERATIVE			AURAL	
	ADV. FORMS*			LOWER A	
	MEAN	(%ILE)		MEAN	(TENTH)
R	46.9	(18)	I	3.8	(5)
V	53.2	(36)	II	4.4	(5)
G	51.2	(43)	III	5.0	(5)
T	50.2	(29)	T	13.2	(5)

*Forms N, O, and P were used in about equal numbers, distributed according to class sections.

Although it is not altogether fair to measure the performance of these students against the one-year Co-operative norm, it is interesting to note that at midyear they were least advanced in Reading, farthest advanced in Grammar, and on the Total were about half-way between the bottom and the median. That their aural scores are in the 5th tenth on all parts of the test is due partly to the fact that they constituted one-third of the cases making up German Norm Group 1A.

The coefficient of correlation between the two tests (total scores) is .46, a passably high figure for a group all phases of whose proficiency in German were in the earliest stages of development.

College G

College G, a large city university, cannot be regarded as producing second-language outcomes typical of the country as a whole. The language students constitute a representative sample of the student body, which as a whole ranks exceptionally high in intelligence as measured by the ACE Psychological Examination. The elementary language course, which meets four hours per week for a total of approximately 105 hours, is designed to give students an adequate reading knowledge of the language in one year. At the same time oral and aural practice is not neglected; the phonetics of the language are studied systematically, and in their four class meetings each week the students have ample practice in handling the spoken idiom in connection with their reading, the bulk of which is covered out of class.

Our testing program at College G consisted in following four successive groups of elementary students in French, and one such group in Spanish. Each group was given the Co-operative tests at 35 hours and at 70 hours, and both the Co-operatives and the aurals at 105 hours, i.e., at the end of the course. The form of the Co-operative test used at each successive testing was the same, and the aural test was the lower level. The following are the results of the final testing.[10]

---

[10] In tabulating the French groups we designate them as Groups 1a, 1b, 1c, and 1d. Group 1c is actually divided into two subgroups, one consisting of students who had had no previous instruction in French, and the other of students who had studied French previously in secondary school.

CO-OPERATIVE			AURAL		
French Group 1a (1944-1945) N=31					
	ADV. R			LOWER A	
	MEAN	(%ILE)		MEAN	(TENTH)
R	84.1	(99)	I	6.4	(7)
V	82.4	(99)	II	16.7	(8)
G	57.6	(65)	III	11.6	(7)
T	77.3	(99)	T	34.7	(8)
French Group 1b (1945-1946) N=13					
R	84.0	(99)	I	8.9	(9)
V	84.1	(99)	II	18.9	(9)
G	54.3	(49)	III	10.9	(5)
T	76.5	(98)	T	38.6	(8)
French Group 1c.1 (Summer 1946) N=13					
R	82.5	(99)	I	6.7	(8)
V	81.0	(99)	II	13.2	(7)
G	58.2	(65)	III	15.0	(8)
T	76.5	(98)	T	34.9	(8)
French Group 1c.2 (Summer 1946) N=15					
R	85.0	(99)	I	7.3	(8)
V	80.5	(98)	II	13.3	(7)
G	59.3	(69)	III	13.4	(7)
T	77.6	(99)	T	34.1	(8)
French Group 1d (Feb.-Aug. 1946) N=16					
R	88.0	(99)	I	10.2	(9)
V	81.8	(99)	II	16.9	(8)
G	56.9	(61)	III	17.3	(8)
T	78.4	(99)	T	44.3	(10)
ADV. C *Spanish Group 1 N=20*					
R	75.3	(99)	I	10.1	(9)
V	79.8	(99)	II	14.6	(9)
G	71.3	(98)	III	18.8	(8)
T	77.8	(99)	T	43.5	(9)

In both French and Spanish, the unfailingly high achievement of these students on the Co-operative tests is due chiefly to the fact that they read exceptionally large amounts in the carefully graded Heath-Chicago series of readers as well as in ungraded material. These readers, like the Co-operative Tests, are based on the frequency counts of vocabulary and structural patterns, and hence the students are well equipped to handle the tests. Doubtless, also, a third exposure to the same form of a Co-operative test is not unlikely to produce a somewhat higher range of scores than would be attained on the third of three different forms. Nevertheless on the first testing, at 35 hours, the percentiles for the total means range from 55 to 84; and on the second testing, at 70 hours, the percentile range is between 91 and 97. Finally it is worth pointing out that performance at the end of *one* year, measured against the *two*-year norms, ranges from the 77th (French Group 1b) to the 89th (Spanish Group 1) percentile.

That, in French, the Grammar scores are consistently far lower than those for Reading and Vocabulary is a direct reflection of lack of stress on grammatical analysis and practice in this course. The Reading and Vocabulary scores are, however, so compensatingly high (actually standing above the 99th percentile) that the total scores still stand at the top. In Spanish there is an equilibrium which reflects a development of grammatical competence on a par with reading ability, which itself runs slightly lower than in French.

Despite the major emphasis on reading, the consistently high aural scores demonstrate that aural skill can be a successful by-product of the "extensive-reading method." Once again the command which these students have of the high frequency vocabulary and constructions stands them in good stead when taking the aural tests based on these same counts. Correlation between the reading and aural skills is generally high as a result, though all the groups are relatively small and hence the coefficients are subject to considerable sampling error. French Group 1d, which made a higher aural score than the others, was made up entirely of war veterans, some of whom had seen service in France or North Africa.

Both French Group 1c and the Spanish group were given the elementary course in a single summer. This was accomplished simply by tripling the pace of the course, with twelve class meetings a week and with no diminution of assigned readings. Judged in the light of the test results, this form of acceleration was neither more nor less effective than the normal pace.

Effect of previous study. It is usual for some students with one or two years, previous experience in the language to enter these elementary courses at College G. Of French Group la, eight of the thirty-one students had already studied French in school. These eight surpassed the rest of the group by 2.0 points on the Co-operative total at the first testing, mainly on the strength of grammatical knowledge; but thereafter, although maintaining a grammatical superiority, they fell slightly below the others on the Co-operative total, and 4.2 points below them on the Aural.

French Group lc was split almost evenly between students with and without high school experience. Differences between their performances at the end of the course are seen to be negligible; at the first testing (35 hours) Group lc.2 surpassed the other by 3.0 points on the Co-operative total, and at the second testing (70 hours) by 2.9 points. At no stage did any one of the differences prove to be significant. As usual, previous study, which does not yield advanced standing, gives a student some initial advantage but this head start is lost by the end of a college year.

High School A

In only one high school, out of dozens that made use of the Investigation's Aural Tests, were the Co-operative Tests administered as well. This was a small private school with a student body selected for intelligence. Foreign languages were taught in small class sections by instructors interested in oral-aural development and competent to teach the spoken idiom, along with reading and grammar, with the maximum of effectiveness possible in a weekly schedule of four hours.

French

The elementary group in French was tested at the end of the year 1945-1946. The second-year and third-year groups we were able to follow throughout the year, giving them the aural tests in October, and both Aurals and Co-operatives in June. The following results were obtained from the three groups:

CO-OPERATIVE AURAL

French Group 1 (N=33)

	JUNE (100 HRS.) ELEM. R			JUNE (100 HRS.) LOWER A	
	MEAN	(%ILE)*		MEAN	(TENTH)
R	49.4	(75)	I	4.9	(7)
V	47.9	(53)	II	9.5	(6)
G	48.5	(52)	III	5.7	(5)
T	48.1	(62)	T	20.1	(6)

French Group 2 (N=27)

	JUNE (200 HRS.) ADV. Q			OCTOBER (110 HRS.) LOWER A			JUNE (200 HRS.) LOWER B
	MEAN	(%ILE)*		MEAN	(TENTH)		MEAN
R	67.6	(84)	I	4.9	(7)	I	12.2
V	64.7	(70)	II	14.8	(8)	II	17.6
G	64.7	(74)	III	9.1	(7)	III	10.0
T	67.2	(77)	T	28.7	(9)	T	39.8

French Group 3 (N=9)

	JUNE (300 HRS.) ADV. Q			OCTOBER (210 HRS.) UPPER A			JUNE (300 HRS.) UPPER B
	MEAN	(%ILE)*		MEAN	(TENTH)		MEAN
R	83.3	(93)	I	11.0	(9)	I	7.2
V	76.9	(73)	II	14.4	(5)	II	16.2
G	73.6	(78)	III	3.4	(6)	III	8.0
T	80.8	(86)	T	28.9	(7)	T	31.4

*The percentiles are from the norms for Public Secondary Schools of the East, Middle West and West: *Booklet of Norms, pp. 21-24.*

It should be kept in mind that the Co-operative norms for high schools are based exclusively on results obtained from public school students. No norms have been issued for private schools; there are reasons for believing that if available, such norms would range higher than the existing ones, and that in competition with other private-school students the present groups might not measure as high as they do. The chances seem good, however, that at least Groups 2 and 3[11] would stand above the median.

These groups' aural scores give evidence of high aural competence as compared with other secondary school students at equivalent stages of study. (The aural norms are based on combined samples from public and private schools, with public schools predominating.) Unfortunately, not enough high-school scores on the B forms of the aural tests were obtained to make possible the establishment of norms. Consequently, aural progress of Groups 2 and 3 from October to June cannot be measured precisely.

Spanish

Only one group was tested in Spanish at High School A, and only at the end of the year. This was a second-year class of 28 students, which at 200 hours produced the following results on the Co-operative and aural tests:

	CO-OPERATIVE ADV. Q			AURAL LOWER A	
	MEAN	(%ILE)*		MEAN	(TENTH)
R	47.3	(24)	I	3.8	(4)
V	46.4	(21)	II	8.1	(4)
G	51.8	(42)	III	12.5	(6)
T	48.6	(30)	T	24.3	(4)

**Booklet of Norms*, pp. 27-30.

It is obvious that the Spanish group's performance was considerably inferior to that of the French second-year class. It seems likely that this difference is attributable to a difference in methods and objectives. The French group read in quantity, aimed to acquire a large high-frequency vocabulary, and covered all the fundamentals of grammar — all of which constituted a certain direct preparation for the Co-operative test; at the same time they received enough ear training to assure them a degree of success with the aural tests, whose word and idiom content they were equipped to handle. The

---

[11] The Group 2 total of 67.2 is at the 43rd percentile of the two-year college norms, and the Group 3 total of 80.8 is at the 68th percentile of the three-year college norms.

Spanish group, on the other hand, was directed toward the complete mastery of a more limited set of objectives. They read comparatively little; they mastered a small vocabulary of a few hundred basic words; they learned to manipulate precisely certain basic grammatical forms (chiefly verbs and pronouns) but did not attempt to cover the full range of Spanish structure; they had relatively little aural training but much work in oral production, which consisted in the memorizing and performing of short skits. Their competence was high within very narrow limits, but it was not the kind that could produce high scores on such broad measures as the Co-operative and aural tests.

SUMMARY

In this chapter we have presented the results obtained by fairly extensive testing programs at seven colleges and one high school. Our purpose has been to give a picture of conventional language teaching, partly to serve as a basis of comparison with the experimental programs we shall examine next and partly to show what the general state of language teaching is.

In regard to the reading skills — to which these schools give primary or almost exclusive attention — the results, on the whole, tend to be better than average. As we have noted, this outcome stems in part from the fact that these are institutions whose reputations for academic standards, calibre of student, and the like are also average or better. We should also expect that schools willing to co-operate with an outside evaluating agency would be those who are proud of their work and convinced that they are doing a better than ordinary job. In the aural skills, to which most of these schools give only very secondary attention, the results run at about the medians of the norms based on conventionally trained students like these.

These generalizations about the programs as a whole should not blind us to the great diversities within them, variations probably produced to some extent by the quality of student and similar factors but undoubtedly also conditioned by the procedures and teaching skill employed. It is difficult to make comparisons fairly; but, whatever the causes of the differences, the reader can readily see how varied are the results produced by "100 hours of conventional college instruction" in French or Spanish.

The numerous correlations drawn from these data (many of which were not reported on the preceding pages because they merely duplicated the established pattern) indicate rather clearly the relation between the reading and aural skills. For all languages, at instructional levels from one college semester to three college years, the vast majority of coefficients between the aural and the Co-operative scores fall within the range .30-.60. (Others, which fall outside this range but within that of .20-.80, are usually definitely explicable for a specific reason and hence should be disregarded.) The usual relation gives little assurance that aural skill necessarily accompanies reading skill or *vice versa*. He who teaches reading will not produce students who necessarily understand the language when spoken; he who teaches aural comprehension will not find his students automatically developing into competent readers. When both the scores in the two sorts of skills and their correlation tend to be high (as, for example, at College G), one finds that a specific reading program has been effectively combined with a specific aural program. This lack of relation between the skills exists even in these conventional programs for whose students the problems of the vocabulary in the aural tests (which we have explained in Chapter 2) does not exist. In the following chapter we shall see how the addition of this further factor makes the cleavage between the skills even greater for the experimental programs.

Chapter VI

Experimental Groups Stressing Aural-Oral Skills as an Initial Objective and Introduction to Reading

Eight colleges and three high schools furnished the Investigation with varying amounts of data — test results and supplementary records — by which to judge the effectiveness of present-day experiments in the teaching of a second language with aural-oral emphasis. We could not hope to investigate any very large percentage of the experiments. We could only select a sample, and this sample was determined by some factors other than the nature of the experiments themselves; for example, by the willingness and ability of course directors to collaborate fully with us in arranging testing programs and compiling supplementary data. By the same token it seems to us probable that the very programs whose directors welcomed and facilitated our investigations were exactly those likely to have the most precise objectives, to be organized the most efficiently, to utilize the most effective methods, and to obtain the most satisfactory results. In other words, even though we were not appraising the complete picture, it seemed unlikely that few if any of its striking features would be absent from that section of it which we were studying.

The programs of our eleven institutions are in agreement in respect to their general objective, which is to turn out students who can use the foreign language in *both* its spoken and its written form. Short of this goal, no two are identical in regard to student selection, type of instructional staff, course organization, teaching materials, special resources employed, and the like. We shall give as full a picture as possible of these particulars as we deal in turn with each experiment; in our concluding chapter we shall make such generalizations as seem justified in the face of the considerable diversity of approach manifested by the several programs.

College H

Our data from College H (a large university with a privately endowed college of arts and sciences) consist of two sets of findings which require separate treatment. A two-year period of limited experimentation at this college preceded a complete reorganization of the foreign-language curriculum with full conversion to semi-intensive oral-aural instruction at the basic level. We were able to follow the original experiment to a certain extent, and the new plan in detail during its first year. While respective results do not lend themselves to significant comparisons, it is interesting to have followed the evolution of experimentation in a single college through two stages, and to note the successive developments in organization and procedure.[1]

1. *Experiments Pre-Dating Full Reorganization of the Language Curriculum*

In 1944-1945 and 1945-1946 the Department of German at College H experimented with new courses at the first- and second-year levels. These courses aimed to develop ability to use the spoken language, and therefore provided for mimicry and memorization of colloquial materials in small drill groups under the guidance of native speakers, as well as for formal instruction in grammar. Reading was also taught, in the latter

[1] Similar evolution of experimentation from tentative beginnings is to be noted elsewhere, e.g.; at College L.

part of the elementary course and throughout the intermediate course. At either level the student was free to enroll in a "regular" group meeting four hours per week (three of drill and one of grammar) for a total of 60 hours per semester, or in an "intensive" group meeting eight hours per week (six of drill and two of grammar) for a total of 120 hours, i.e., the equivalent of one year in a single semester.

The Investigation's tests were administered at irregular intervals to both intensive and regular groups during 1944-1945 and 1945-1946, but the resultant data are so diffuse that no meaningful conclusions can be drawn from them. For one thing, some students took the *intensive elementary* course and afterward continued with the *regular intermediate* course, while others who had taken the *regular elementary* course followed it with the *intensive intermediate* course. This brought about a shifting of group enrollment from semester to semester, so that no single group remained homogeneous; and since testing was casual, few individual students were followed straight through two years' work.[2] All in all, the results are not conclusive enough to warrant a statement here concerning the effectiveness of these first experiments in German at College H.

2. *Reorganization of the Foreign-Language Curriculum*

In the fall of 1946, a completely revised program of modern foreign-language teaching was instituted at College H. Administratively the new plan brought basic-level instruction in Romance, Germanic, Slavic, and Oriental languages together into a single modern-language division, with unified objectives and procedures.[3] Educationally, the revision constituted a civilian adaptation of the type of program originally planned for the ASTP, and it was an adaptation made on the largest scale yet attempted in an American college, affecting all elementary and intermediate language students in the institution. Successful functioning of the plan during the first years of experimentation and adjustment was insured by a substantial grant to the university from a national foundation.

---

[2] Test scores for 1945-1946 are tabulated in Appendix B.

[3] Chinese, French, German, Italian, Russian, and Spanish were offered in 1946-1947. The addition of Portugese was scheduled for 1947-1948, and of Polish for 1948-1949.

We believe that the "orthodoxy" of the new plan, from the point of view of the kind of oral-aural approach developed in the ACLS Intensive Language Program (see above, p. 1) and in the ASTP, justifies the devotion of several paragraphs to its objectives, organization and procedures. In respect not only to this orthodoxy itself, but to the extent of its application within a single institution, the College H plan is our most outstanding example of extensive experimentation.

Objectives. The new plan aims to develop basic competence in (1) speaking and understanding the language as used by natives, and (2) in reading works written in the language. Both these abilities are regarded as desirable and legitimate goals within the field of liberal education. Given this two-fold objective, the assumption underlying the method of instruction is that the most efficient and most thorough way to comprehensive mastery of a language is through initial preoccupation with its spoken form as used by natives in their daily communications. Thus, learning to use the spoken language becomes the initial aim of instruction, toward which a major effort is directed in the first semester of study.

Organization and procedure. In order to provide the continuous opportunity for practice necessary to the development of oral and aural facility, the elementary and intermediate courses are organized on a semi-intensive basis of eight contact hours per week. Within this eight-hour system the two-fold approach co-ordinating analysis and drill practice is employed. Two of the eight hours are spent in a large lecture section, accommodating up to fifty students, presided over by the linguist in charge of the course and devoted to discussions on the structure of the language. The remaining six contact hours are spent in drill sections devoted to oral and aural practice (and ultimately also to work with reading texts) under the supervision of a native-speaking instructor; in these sessions the opportunity for maximum individual participation is insured by limiting the number of students to ten.

A basic sequence of two courses is provided: the elementary and the intermediate. Each covers one fifteen-week semester and thus at eight hours per week accumulates a year's total of 240 contact hours. Outside preparation for class hours is expected in approximately a one-to-one ratio; course credits assigned are six per semester as contrasted with three for other courses; thus all in all, the one-semester intensive course is to be equated with a one-year course, and the two-semester sequence with a two-year course.

The staff participating in the program consists of a ranking member in charge of each language, whose role is *(a)* to teach the large lecture sections, endeavoring to apply modern techniques of linguistic analysis to the material being concurrently assimilated by the learners in the drill sessions, and *(b)* to organize and supervise the work of the native instructors and to co-ordinate the activity of the numerous drill sections. The natives who have instructorial or graduate-assistant status are recruited from a variety of sources: a few were members of the departmental staff before the adoption of the new plan; others are natives expressly imported to serve on a full-time teaching basis; several are foreign graduate students pursuing their own studies at the university and earning tuition by serving as language teachers. In the future it is planned to procure a steady supply of such personnel on a basis of mutual exchange of students between College H and foreign institutions. It goes without saying that, in such a plan of staff procurement, experience in teaching one's native language is not a prerequisite for service. The plan assumes that success in drilling students orally and aurally in one's own language does not require pedagogical training of the conventional sort, and that such training may indeed lead an instructor to deviate from his primary function and attempt to teach the grammar of his language according to criteria other than those adhered to by the linguist handling the analytical phase of the course. Instructors are transferred from one drill group to another at least once during each semester, so that a student's experience with the spoken language includes familiarity with more than one native speech-pattern. As soon as the individual students' learning ability and progress rate can be reliably appraised — normally after six or eight weeks of instruction, — they are segregated into "fast" and "normal" groups in order that all may benefit by drill techniques best adapted to their capabilities.

The texts used for oral-aural work in the elementary and intermediate courses are the series entitled *Spoken French, Spoken German,* etc. originally prepared for the United States Armed Forces Institute and now available for general civilian use. All follow essentially the same plan and are well suited to use in an eight-hour-per-week program: each lesson, or *unit,* contains a sufficient amount of drill material for six hours,[4] and a "word study" section around which the linguist may profitably organize two hours of discussion. Approximately half of the thirty units are covered each semester.

After instruction in reading is begun, late in the elementary course or at the beginning of the intermediate, two or three of the six weekly drill hours are devoted to it. While explanation of puzzling words or constructions is provided, translation of the text into English is avoided in

[4] For example: mimicry-memorization of "Basic Sentences", 2-3 hours; aural practice with the "Listening In" section, 1 hour; controlled oral practice ("What would you say?"), 1 hour; free conversation, 1-2 hours.

favor of oral reading, paraphrasing, and question-and-answer based on the content. With the possible exception of some dictation from the reading-text, no writing in the language is done by the students throughout the sequence. On the assumption that writing ability is logically the last and most difficult of the skills to develop, it is taught only in the advanced courses.

Audio-visual aids were not used extensively during the first year of the program, for two reasons: (1) it was not a good year for the purchase of electronic recorders and similar devices, as technical improvements were being made so rapidly that successive makes of machines were quickly becoming obsolete; (2) dispersion of class sessions throughout many buildings of the campus (a situation which was to be remedied by the fall of 1947) made even the use of phonographs difficult. However, toward the end of the year listening-rooms were set up and many students availed themselves of the opportunity to listen (on their own time) to the recordings of the basic text series, and in German to some supplementary recordings cut by the local staff.

The foreign-language requirement. Along with the reorganization of instruction, the language requirement for graduation[5] was redefined in terms of proficiency rather than accrued course credits — in the words of the director, of "achievement rather than endurance". The level of achievement which actually constitutes proficiency is determined by the language staff on the basis of performance on standardized and other examinations. Theoretically any student may present himself for the proficiency examination at any time it is administered, whether or not he has completed the basic sequence of courses. In practice, such student homogeneity prevails as regards language experience and competence that the bulk of the students present themselves for proficiency at the end of the basic sequence, and at that time the majority of them are declared proficient.

---

[5] This requirement may be fulfilled in any one of the languages taught.

Placement. The bulk of the elementary-course enrollment naturally consists of students having no previous experience with the language in question. Students who have studied the language in high school take a placement examination which is identical with the examination used to determine proficiency. On the basis of their scores on this placement test they are (1) required to start in the elementary course[6], (2) placed in the intermediate course, or (3) declared already proficient and permitted to enroll in an advanced course if they wish. A student placed in the elementary course may take the proficiency test again at the end of that course, or he may wait until he has completed the two-course sequence, the point at which most students take the proficiency examination. If at this stage he fails to perform at the proficiency level but still passes the course, he is required to study the language for one more semester, in a three-hour-per-week course designed chiefly to bring his reading ability up to an acceptable level of achievement. He is required merely to pass this course, not to retake and pass the proficiency examination. It is felt that this arrangement encourages striving for proficiency because it rewards the superior student by early satisfaction of his language requirement and penalizes the dull or lazy student by requiring extra work.

The desirability of using standardized measures for placement and proficiency determination led to the adoption, for this purpose, of the Co-operative Tests and the Investigation's Upper Level Aural Comprehension Tests. The resulting data was doubly useful, serving not only the present purposes of the Investigation, but also fulfilling an integral function within the new language program at College H.

---

[6] In this situation they are not deprived of their entrance credit in the language.

Testing data. Since each language was taught in accordance with the same principles and procedures, the presentation can be made according to major course-groups each comprising students of the four languages in which our tests were available, namely, French, German, Russian, and Spanish. The subdivision of elementary students into Groups 1a and 1b, and that of intermediate students into Groups 2a and 2b, will be made clear as we proceed. Allotment of hours to the different phases of instruction was as follows:

		ELEMENTARY	INTER-MEDIATE	SEQUENCE TOTAL
Oral-Aural Drill	French	75	64	139
	German	75	45	120
	Russian	90	60	150
	Spanish	65	60	125
Reading*	French	0	26	26
	German	0	45	45
	Russian	0	30	30
	Spanish	10	30	40
Analysis Lectures	French, German, Spanish	45	30	75
	Russian	30	30	60

Group 1a. These students took the elementary course in the first term of 1946-1947 and the intermediate course in the second term. The numbers completing the two-year sequence were: French, 46; German, 113; Russian, 19; Spanish, 70[7].

*The following texts were read: French: Parker and Grigaut, *Initiation à la culture française;* German: Kästner, *Emil und die Detektive* and Gaede and Klug, *Geschichten von Heute und Gestern* (by all students); Röseler, *Moderne Deutsche Erzähler* (by one third of the students); Russian: Pargment, *Elementary Conversational Russian Reader;* Spanish: Del Río and Hespelt, *Lecturas Hispánicas, Book I* and Fernández, *Por Esas Españas.*

[7] Enrollment in the first term was larger by about 15%. The attrition resulted from three causes: (1) returning veterans were permitted to waive certain required courses, and some chose to waive completion of the language requirement; (2) students who had completed their language requirement in some other language were free to discontinue the present one; (3) some students from other colleges than Arts and Sciences, having no foreign language requirement, also discontinued. Our data for Group 1a is based exclusively on the continuing cases.

The results of the Co-operative and Aural test administrations in February and in June were as follows:

	CO-OPERATIVE					AURAL			
				French (N=46)					
	FEB. (1 YR.) ADV. R		JUNE (2 YRS.) ADV. Q			FEB. (1 YR.) LOWER A		JUNE (2 YRS.) UPPER A	
	MEAN	(%ILE)	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH
R	52.3	(41)	69.1	(57)	I	8.0	(9)	12.0	(7)
V	59.1	(48)	73.5	(55)	II	15.1	(8)	17.8	(5)
G	56.4	(59)	63.9	(61)	III	9.6	(5)	3.2	(6)
T	56.5	(49)	70.6	(58)	T	32.8	(8)	33.0	(7)
				German (N=113)					
	ADV. Q		ADV. P						
R	49.9	(27)	58.7	(21)	I	6.9	(6)	6.5	(5)
V	43.2	(8)	59.5	(21)	II	5.3	(4)	17.4	(6)
G	55.5	(61)	58.4	(33)	III	7.8	(5)	5.9	(5)
T	49.8	(29)	59.7	(22)	T	20.0	(5)	29.8	(8)
				Spanish (N=70)					
	ADV. Q		ADV. P						
R	55.3	(56)	66.1	(51)	I	7.8	(8)	10.9	(7)
V	52.8	(38)	60.7	(27)	II	12.7	(7)	25.3	(7)
G	50.6	(42)	64.0	(56)	III	14.6	(6)	3.8	(6)
T	53.7	(46)	65.0	(44)	T	35.1	(7)	40.0	(7)
				Russian (N=19)					
	INVESTIGATION'S TEST	(TENTH)		(TENTH)					
R	5.2	(1)	20.5	(6)	I	5.2	(3)	6.6	(3)
V	5.2	(1)	16.6	(7)	II	11.1	(4)	17.5	(4)
G	12.9	(4)	21.1	(7)	III	14.9	(4)	5.4	(4)
T	23.3	(1)	58.2	(6)	T	31.2	(4)	29.5	(3)

Observations on Co-operative scores. In February, the means of the total-scores for all groups were sub-median. By the end of the year those for French and Russian reached or exceeded the two-year norms, but those for German and Spanish fell short. The averages for Reading and Vocabulary show the same general trend. Grammar tended to be the strong point for all groups on both administrations.

Observations on aural scores. In French and Spanish, which according to the lower-level norms were more successful than German and Russian at the elementary stage, there is a tendency toward greater strength on Parts I and III (short items) than on Part III (anecdotes). All the aural practice material in the first half of the spoken language texts consists of short utterances, while only in the second half are a few anecdotes introduced.

The comparative lowness of the Russian scores is due in part to the nature of the norm-distributions, which included (*a*) some groups of students more advanced than these, and (*b*) many students with Russian or related Slavic backgrounds.

Previous experience of Group 1a students with the language studied. Eight French, ten German and eight Spanish members of Group 1a entered the elementary course in the fall having studied the particular language in high school, some of them as much as two years. They accordingly took the placement examination in October,[8] and the experience of coping with forms of the Co-operative and aural tests probably added to the advantage they enjoyed over the raw beginners when all were tested in February. The following tabulation shows the differences (between the means of the total scores) by which the experienced students surpassed the inexperienced on the Co-operative and aural tests in February and in June:

	CO-OPERATIVE		AURAL	
	FEBRUARY	JUNE	FEBRUARY	JUNE
French	14.6	1.8	6.3	2.4
German	8.8	4.9	4.9	6.0
Spanish	12.6	7.1	14.4	5.8

In the reading skills, the students with previous experience were solidly enough grounded to maintain an apparent advantage over the novices throughout the two-year sequence in all languages, but none of the differences is significant by

---

[8] Their scores are given in Appendix B.

June, partly (though probably not wholly) because the subgroup then took Upper Level Form A for the second time, the first time having been on placement in October. But for both types of test, the numbers of students are too small for the results to be significant. In the field of oral production, on the other hand, although we have no objective data, it is reported that many of the students with previous experience remained inferior in pronunciation (though not necessarily in fluency) to many of the raw beginners whose speech habits were first formed through direct imitation of native speakers.

Group 1b. These students took the first-year course in the second term of 1946-1947 and were due to complete the sequence in the first term of 1947-1948. The numbers: French, 41; German, 106; Spanish, 61.[9]

	CO-OPERATIVE			AURAL	
French (N=41)					
	JUNE (1 YR.)			JUNE (1 YR.)	
	ADV. R			LOWER B	
	MEAN	(%ILE)		MEAN	(TENTH)
R	57.8	(62)	I	8.4	(7)
V	60.9	(56)	II	13.0	(5)
G	58.7	(68)	III	7.9	(5)
T	60.8	(67)	T	29.4	(7)
German (N=106)					
	ADV. Q			LOWER A	
R	49.7	(29)	I	6.8	(6)
V	44.2	(10)	II	5.0	(4)
G	57.1	(67)	III	6.9	(4)
T	50.2	(30)	T	18.7	(5)
Spanish (N=61)					
	ADV. Q.			LOWER B	
R	53.2	(47)	I	10.0	(8)
V	50.3	(30)	II	9.8	(6)
G	45.3	(23)	III	9.1	(2)
T	49.5	(30)	T	28.9	(5)

---

[9] Elementary Russian was not given in the "off" term.

Observations comparing these Group 1b results with those of Group 1a at the end of one year:

1. French results were slightly better on all parts of the Co-operative test and slightly poorer in aural comprehension.

2. German Co-operative scores were approximately the same in each area (high only in Grammar); aural results were likewise the same.

3. Spanish scores were distinctly poorer in Grammar, and in Part III (Anecdotes) of the aural test.

Although the directors were striving to effect improvements on the basis of their experience, it is clear that the results attained did not correspond to their efforts. At least three considerations seem to affect this situation: (1) the interest and energy of tiring drill instructors tended to be directed more toward assuring the continuing success of the original Group 1a (then in its intermediate stage) than toward perfecting a second group of beginners in the same year; (2) fewer students in Group 1b had studied the language previously, whereas in Group 1a such students contributed measurably to raising the average of the group as a whole; (3) Group 1b contained some students for whom the elementary language course was mainly a six-credit filler in their second-term schedules; they did not propose to continue the language another year, and they were satisfied to meet minimum standards.[10]

Group 2a. These students took the second-year course in the first term of 1946-1947. The group was made up of (1) entering freshmen with an average of three years' study of the language in high school, and (2) students who had already successfully completed one year of the language at College H.[11]

[10] Frivolous elections of an elementary language course are more likely to occur in German or Spanish -- often as the *second* foreign language studied -- than in French, the generally commonest *first* foreign language.

[11] For some, this "one year" had consisted of a concentrated elementary course of 90 hours during the previous summer at the rate of fourteen hours per week.

The results of the tests are given below, together with those of Group 2b.

Group 2b. This small group of students took the second-year course in the second term of 1946-1947. In class they were amalgamated with Group 1a, from whom, however, they differed in that they had not taken the first-year course along with them. Their background was identical with that of Group 2a, described above. The results of testing Groups 2a and 2b are as follows:

	CO-OPERATIVE			AURAL	
	French 2a (N=95)				
	FEB. (2 YRS.) ADV. Q			FEB. (2 YRS.) UPPER B	
	MEAN	(%ILE)		MEAN	(TENTH)
R	69.8	(60)	I	7.9	(8)
V	75.3	(62)	II	18.3	(9)
G	66.9	(72)	III	7.3	(8)
T	72.9	(66)	T	33.6	(9)
	German 2a (N=98)				
	ADV. P			UPPER A	
R	62.3	(32)	I	5.5	(5)
V	62.2	(29)	II	18.7	(7)
G	61.6	(47)	III	6.4	(6)
T	63.1	(34)	T	30.6	(8)
	Spanish 2a (N=105)				
				UPPER B	
R	71.2	(71)	I	11.7	(7)
V	68.8	(58)	II	9.1	(6)
G	70.5	(79)	III	5.4	(4)
T	72.3	(72)	T	26.2	(7)
	French 2b (N=16)				
	JUNE (2 YRS.) ADV. O			JUNE (2 YRS.) UPPER A	
R	77.6	(84)	I	13.7	(8)
V	79.1	(76)	II	20.8	(7)
G	71.0	(86)	III	5.0	(8)
T	78.7	(85)	T	39.5	(8)
	German 2b (N=24)				
	ADV. P				
R	62.7	(34)	I	6.3	(5)
V	63.5	(34)	II	16.3	(6)
G	60.8	(44)	III	6.0	(5)
T	63.5	(35)	T	28.6	(7)

Spanish 2b (N=10)

R	70.8	(70)	I	11.7	(7)
V	65.1	(42)	II	24.7	(7)
G	69.2	(76)	III	2.4	(3)
T	70.4	(67)	T	38.8	(7)

In comparing the achievement of Groups 2a and 2b (students who entered the basic sequence at the intermediate level after previous study of the language elsewhere) with that of Group 1a (those who took the basic sequence from the beginning), we see:

1. In French, Groups 2a-2b were strikingly better on all the Co-operative scores and generally better on aural parts and total.

2. In German, Group 2a-2b's Co-operative scores were generally higher, though still sub-median in all areas; aural results differed little.

3. In Spanish, the Co-operative scores of Group 2a-2b were decidedly better in every area, while in aural scores there was no difference.

High-school versus college experience at the elementary level in Group 2a. Thirty French, eighteen German, and twenty-four Spanish members of Group 2a entered the intermediate course in the fall having studied the particular language in high school. They accordingly took the placement examination in October[12] and may have been thereby prepared somewhat for facing the Co-operative and aural tests again at the end of the term. The following tabulation shows the difference in total means by which the high-school-trained surpassed the college-trained subgroup on the Co-operative and aural tests at the end of the intermediate course in February:

---

[12] Their scores on this test appear in Appendix B, as do those of the few members of Group 2b who took the placement examination in February. Students entering from a course given previously at the college itself were not required to take the placement examination in October; a handful did take it in February.

	CO-OPERATIVE	AURAL
French	5.2	2.1
German	9.8	16.3
Spanish	10.3	5.7

The high-school subgroup's preparation for the intermediate course seems to have been superior in both the visual and the aural area. (In German, the high aural differential is partly due to the fact that the same form of the upper-level test was used both for the placement and in February.)

Comparison with Conventional Groups

A number of upper-class students at College H completed their language requirement by taking a conventional intermediate course given for the last time in the fall term of 1946-1947. The principal aims of the course, which met three hours per week, were reading ability and a review of grammar, with little stress on aural-oral achievement. Twenty-five French, forty-five German, and forty-three Spanish students taking the conventional course were tested by the Investigation in February. Their scores will be found in Appendix B. The following table lists the French, German, and Spanish Groups 2a described above, which were completing the intermediate course under the new plan at the same time. In the Co-operative and Aural columns appear the differences in means between these Groups 2a and their conventional counterparts. A plus sign means that the difference favors Group 2a; a minus sign means that it favors the conventional group.

Differences between New-Plan Group 2a and Conventional Groups

CO-OPERATIVE			AURAL	
	2 YRS.	*French*		2 YRS.
R	+7.7		I	+1.5
V	+6.1		II	+9.0
G	+4.5		III	+3.6
T	+6.7		T	+14.2
		German		
R	+2.5		I	+0.2
V	+0.8		II	+5.8
G	+6.3		III	+0.2
T	+3.4		T	+6.2
		Spanish		
R	+1.0		I	+1.5
V	-1.8		II	+1.6
G	-0.2		III	+0.1
T	-0.2		T	+3.2

While the differences mostly favor Group 2a, little weight can be attached to them because the conventional group constituted a heterogeneous mixture of irregular cases with varying amounts of previous experience and generally poor linguistic aptitude.

Correlations

Correlations were computed for all intensive groups taking (other than for placement) Lower or Upper Level Form A of the aural together with the Co-operative Test, in either February or June. The French and Spanish groups took Aural Form B in June, and are therefore excluded from the following figures:

Lower-Level Aural and Co-operative Totals[13]	
French (Group 1a, February):	.59
German (Group 1a, Feb., and 1b, June):	.50
Spanish (Group 1a, February):	.51

Upper-Level Aural and Co-operative Totals[13]	
French (Groups 1a and 2b, June):	.64
German (Groups 2a, Feb., and 1a and 2b, June):	.40
Spanish (Groups 1a and 2b, June):	.16

---

[13] Correlations of part scores on these tests will be found in Appendix C.

These figures indicate that in experimental intensive courses aiming to develop both aural and visual skills on the assumption that the first contribute to the excellence of the second, correlations between outcomes are no higher than in conventional courses in which reading is the chief objective and aural comprehension a secondary aim. Further data corroborating this observation will be encountered in our study of other experimental programs (e.g., College J).

Summary

Fundamental problems of teaching were not solved, nor crucial issues of learning settled, in the first year's operation of the new plan at College H. On the basis of the 1946-1947 data alone we are not justified in appraising its present effectiveness and future potentialities. The first steps of an operation as vast and complex as this were inevitably plagued by handicaps of every description. The corps of native-speaking tutors was new and untried. With enrollment figures in the new scheme unpredictable up to the last minute, it was impossible to estimate the exact number of tutors needed, and as a result a few had to be hastily recruited from among native speakers available on the campus. In the future, enrollment figures will be known in advance and staff needs more readily estimated. Efficient instructors will be retained, and will resume work fortified by their past experience; time will permit a considered selection of replacements for those who were unsuccessful. During the first year a shortage of classroom space necessitated the assigning of offices and the scheduling of lectures and drill sections at widely scattered points on the campus. In 1947-1948 the entire foreign-language plant will be housed in a single building, which will facilitate (*a*) intercommunication between course directors and instructors, (*b*) co-ordination and improvement of classroom procedure through frequent observational visits by course directors, (*c*) use of audio-visual equipment.

The gradual overcoming of physical handicaps, with an attendant increase in efficiency of operation, should permit the staff to concentrate on improvements of teaching technique leading directly to levels of performance higher than those reached during the first year. They know that in order to achieve that objective of high performance they need to work constantly to effect improvements at the classroom level, and they know that improvements at the classroom level mean the selection and adoption (or the creation, if need be) of the finest possible textual materials, the effective utilization of audio-visual aids, and — most important of all — the continuing in-service training of teachers.

College J

In the fall of 1944 College J (a large privately endowed university) instituted a new Bachelor of Arts program characterized mainly by a fixed curriculum. The number of freshman students embarking upon this program was relatively small as compared to the total arts-college enrollment; it was therefore possible to establish, without great increases in staff, a semi-intensive experimental course in foreign language (French, German, or Spanish) for the entire new B.A. group. All language courses outside the B.A. program remained of the conventional type, and could thus serve as a control against which to measure the outcomes from the experimental courses.

The experimental courses at College J roughly doubled the conventional amount of four contact hours per week prevailing in the basic-level courses. It was not intended, however, that the students should thereby reach or surpass the customary levels of performance in half as much time as the conventional groups. The primary justification of the increased contact rate was the notion that the traditional "home work" required of the students outside of class should be supplanted by in-class drill work, i.e., frequent supervised practice in using the spoken language with the purpose of developing oral and aural skills. Outside preparation for the drill itself was generally not expected, and consequently the total amount of time devoted to the language was, theoretically at least, much the same in experimental as in conventional courses.

In view of the very different natures of the two curricula in which each language program was offered, it seemed possible that the quality of student might also be different. For example, the B.A. curriculum might tend to appeal to students of superior interest and ability in foreign languages, and such selective factors would, of course, invalidate comparisons between the achievements of the two language programs. Unfortunately, because of the heterogeneity of the B.S. curriculum and consequent scattering of the students' records through all the many divisions and schools of the university, it was impossible to secure the data for exact comparison of each experimental group with its conventional (B.S.) counterpart. Complete data for the B.A. students were obtainable. For the B.S. group, a random sample of equal size with the same representation of the languages was selected.

Two bases of comparison were available: intelligence (as measured by the California Mental Abilities Test) and facility in the native language (as reflected by the grade in Freshman English, the only course common to students in both curricula.

For 1944-1945 the observable differences on both these scales favored the B.A. students, though not significantly. In 1945-1946 differences in the same direction were significant at the 5% level for the Mental Abilities Test but not for the English grade.

Another possible index seemed to be the grades assigned to students in the different programs by the particular departments. In French during 1944-1945 the grades awarded to the B.S. sections were slightly higher, though not significantly so, than those given to the B.A. students; in 1945-1946 the equally slight trend was in the other direction. The German Department awarded equal grades to both groups in 1944-1945, but in 1945-1946 it favored the B.A. students to the degree of the 5% level of significance. The Spanish Department favored the B.A. students slightly both years, but in neither case was the difference significant.

For both groups in both years the mutual correlations between these three criteria (Mental Ability Score, English grade, and foreign language grade) are strikingly similar, with .38 being the median coefficient for most combinations.

In short, the available evidence suggests that while the B.A. students in the experimental language program may have been slightly superior in 1945-1946, the differences are very small and that the groups compared in the following pages are probably equivalent in intelligence and linguistic facility.

The Investigation followed all first, second, and third year language courses, both experimental and conventional, through 1944-1945 in two languages and through 1945-1946 in three languages; and experimental groups in two languages through 1946-1947. All groups were tested with the Co-operative tests at the end of each quarter and were given the aural tests twice each year. In this chapter we report only the results obtained when both tests were administered simultaneously; the scores on the full battery of tests appear in Appendix B.

We present the data by languages, as procedures varied somewhat from one to another.

French

Group 1a. Seventeen B.A. students took the elementary experimental course in 1944-1945, and nine of them continued with the intermediate course the following year. In the first quarter of the first year, the language portion of their curriculum was devoted to a general course in linguistics (entitled "The Nature of Language"), conceived as preparatory ground-work for beginning the study of their particular second language in the second quarter (January). During the second and third quarters they met for lecture and drill ten hours per week for a total of 180 contact hours.[14] Their principal textbook was Denoeu and Hall's *Spoken French,* of the Heath Spoken Language series. Prescribed procedure in the use of this manual was followed: i.e., lectures dealt with the phonology, morphology, and syntax of French (with examples drawn from the material already assimilated) and with the phonograph records which reproduce the text utilized for illustration; drill work was devoted to mimicry and memorization of the basic sentences, aural practice listening to variations of the basic matter, oral exercises of various types, and finally conversation practice in reproducing the situations treated. The lectures were given by the course director to the full class meeting together; for drill the students were divided into two groups, and the sessions were conducted by native or near-native instructors all of whom were regular members of the department staff and had had some classroom experience in teaching French.

During the three quarters of the second year, Group 1a met eight hours per week[15] and accumulated a two-year total of 390 hours. They worked principally with the text of a contemporary French play, *Les Jours Heureux*. The previous year, the director of the course had prepared a review analysis of French grammar, organized by structural categories but based exclusively on patterns occurring in the text and already familiar to the students when treated; this analysis, supplemented with exercises, constituted the material of the lectures. In the drill sessions the group read the play aloud, held conversations on its content, and dramatized certain scenes with roles assigned and lines delivered from memory.[16]

---

[14] Four hours per week were devoted to lecture, six to drill.

[15] Four in lecture and four in drill.

The nine students in Group 1a who completed two years of French[17] produced the following test scores:

CO-OPERATIVE

	JUNE 1945		DEC. 1945	JUNE 1946	
	1 YR. ADV. O		1 YR. & 1 QU. ADV. R	2 YRS. ADV. R	
	MEAN	(%ILE)	MEAN	MEAN	(%ILE)
R	41.6	(10)	52.4	65.8	(45)
V	63.8	(67)	60.4	66.1	(25)
G	55.0	(53)	57.3	61.6	(50)
T	53.8	(38)	57.2	66.0	(39)

AURAL

	JUNE 1945		DEC. 1945		JUNE 1946	
	1 YR. LOWER A		1 YR. & 1 QU. UPPER B		2 YRS. UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)	MEAN	(TENTH)
I	6.9	(8)	2.0	(2)	11.3	(6)
II	12.9	(7)	11.1	(5)	18.2	(5)
III	15.3	(8)	7.3	(8)	3.9	(7)
T	35.1	(8)	20.4	(5)	33.4	(7)

The Co-operative scores indicate that the group (*a*) did not learn to read at all adequately in one year and was still below the median at two years, (*b*) developed very respectable vocabulary for the one-year level but failed to increase that Vocabulary score very much in the second year, (*c*) acquired an average knowledge of grammar.

The aural scores show (*a*) very good performances at the end of one year, (*b*) average comprehension of the difficult upper-level material after only one quarter of the second year, particularly the dialogue section (taken, it will be recalled, from a play), and (*c*) a continuous development of aural ability through the second year.

---

[16] As will be reported below, the dramatization principle rather casually exploited in this group became the mainspring of an experiment with a subsequent group of French beginners.

[17] See Appendix B for scores from the original seventeen at one year.

Group 1b. Inspired by his promising experiments in developing oral facility in French through the classroom use of a French play, the director conceived the plan of an introductory course which would exploit the full resources of dramatics in the teaching of French. He believed that the common human liking for play-acting would lend motivation to some students in whom otherwise it might be lacking, and that the sense of taking part in a dramatic performance would bring some release from the self-consciousness which seems to inhibit many individuals' attempts at oral production in a foreign language. He felt, however, that to be conclusive the experiment must be carried out under certain optimum conditions, which included (*a*) the use of a theater with a stage on which the students could act their productions, (*b*) some help in the techniques of staging, (*c*) specially prepared materials. Our colleague's search for these requisites was satisfactorily rewarded: (*a*) the university partially financed the conversion of a campus attic into a small theater; (*b*) a grant from a national foundation provided the aid of a dramatics coach as well as meeting the balance of the construction costs; (*c*) a member of the Investigation staff undertook to collaborate with the director in the production of the teaching materials. The course was inaugurated in the fall quarter of 1945.

Twenty-five students were enrolled. Seven hours per week of their time was at the teaching staff's disposal.[18] The normal distribution of these seven hours was three for lecture and four for drill.

The basic materials consisted of a series of eleven original skits, written by the course director. The recurring *dramatis personae* of the successive playlets are an American family visiting in Paris, and a French family who are their special friends and hosts. Each scene embodies some situation of French life as it might be experienced by foreign visitors. Thus in the first two scenes the Americans are greeted by their friends at the boat and go through the customs; in subsequent episodes they travel by train, find hotel accommodations, eat at a restaurant, spend Christmas eve in their friends' home; the younger generation goes to the hairdresser's, partakes of French social life, attends a *lycée*. The vocabulary of the skits is not wholly utilitarian; many specific terms occur which, it is true, are rarely heard outside of restaurants or beauty parlors; but also many common words and expressions are used by the characters as they converse which are transferable to almost any life situation. No concession is made to simplification of vocabulary or idiom. The style is intended to be realistic, the tone sprightly and sophisticated; expressions occur which are vividly colloquial, some of them bordering on slang.

---

[18] For the year 1945-1946, the B.A. curriculum was revised in such a way that elementary language students began their particular language in the first quarter and met seven hours each week throughout the year, with an eighth hour per week devoted to the general linguistics course which in the previous year had been completed in the first quarter.

The text of each skit was mimeographed in two columns, with an equally colloquial non-literal English version paralleling the French. The French version was not printed in normal French orthography at all: in each unit it appeared only in the form of a phonemic transcription, of which we reproduce a short sample from one skit:

Boy: vwasil trãtéũ, médam. aprè vou, silvouplé.	Here's number 31, ladies. After you please.
Mrs. R: sè-tünbèl chã̲br, asé grã̲d/é trè klè̲r.	It's a nice room— quite large and very light.
Boy: wi, madam, é vouzavé ló kourã̲t, chó-dé frwad, ó lavabó isi.	Yes Ma'am; and you have running water, hot and cold, in this lavatory here.
Marie: kèl ròbinè pourló chó̲d, é l^kèl pourló frwad?	Which faucet is for the hot water, and which is for the cold?
Boy: j^krwak-sè s^lẅi agó̲ch/pourló frwad, mé sè peu-tè̲tr/adrwat.	(hesitating) I think it's the left one for cold, but maybe it's the right one.

This transcription was deliberately devised to apply and test the theory that learners of French will acquire an authentic pronunciation of French if they are not too early confronted with its written form and thereby influenced by appearances before their habits of sound-production are fully formed. Such symbols as the *ch* and *j* were adopted in order to minimize the new steps involved later in approaching the regular spelling. Students using the Heath *Spoken French* deal first with phonemic transcription but their transfer to normal spelling is begun after about 35 hours. The present group, on the other hand, saw no conventional written French whatever as long as they were learning and dramatizing their playlets, that is to say until about the middle of their third quarter. In preparing each skit they first heard the material spoken by their instructors and then imitated that spoken model; the transcription served as a useful and (for most persons) necessary visual aid during the basic aural process, and could later be consulted by the student as a completely reliable guide to the sequences of sounds, which it represented in one-to-one correlation with its symbols.

The procedure employed in working with the playlets was as follows. The group were brought together in a two-hour lecture period. Before they received their typed copies of the unit, their dramatics coach summarized its content in English. Next the staff of five instructors, augmented if necessary by one or two pre-trained members of the student group, presented the skit with full action and "props" on the stage while the students merely looked and listened. They then received the printed text; the coach explained the "business", some of which they noted down in the margin of their copy; the scene was then repeated by the staff with the students following the text. The group then divided into four drill sections to prepare for their own presentation of the skit. Two weeks of drill, totaling eight hours, were spent in learning and rehearsing a unit: six hours perfecting the lines, and two hours walking through the parts under the guidance of the coach. The drill practice consisted of mimicry-memorization, choral at first, and later individual as each student concentrated on his assigned role. The culmination of two weeks' work was of course the final presentation of the skit by each of the four drill sections.

There is no doubt that the students enjoyed their "French in action" much more than a language course is normally enjoyed. Cutting classes was extremely rare, and every member of the group worked hard to perform at maximum capacity. Even as new roles were learned every two weeks, earlier ones were remembered without difficulty and previous skits were frequently replayed on special occasions. With the exception of one or two students who could neither imitate nor memorize the French lines and had to be dropped from the course, all members of the group acquired acceptable pronunciation and intonation, and the casual spectator at one of their presentations received the impression that these students really knew French. Actually, the French that they knew was limited strictly to the material of the skits. They could neither understand nor utter a single expression which did not occur verbatim somewhere within their script. Therein lay the weakness of the method, which could have been supplemented more fully than it was.

On the theory that the students should understand the structure of the phrases they were uttering on the stage, a descriptive analysis of the material was issued to them in blocks covering several units at a time. This treatment was based strictly on the grammar of spoken French, in order that all forms or combinations of forms discussed could be illustrated in terms of the phonemic transcription. Supplementary exercises were provided in the form of short English expressions for which the French equivalent had to be arrived at analytically. Chiefly because of the relatively little class time devoted to this (two or three lecture hours at widely spaced intervals), the students failed to assimilate this new and unfamiliar kind of material satisfactorily.

In addition to the grammatical analysis, shorter English versions of each skit were provided which introduced variations or twists not encountered in the original. These were designed to be rendered into French, without memorization or staging, for practice in varying the patterns of the principal versions. For the earlier units these variations involved mere transposition of set phrases without internal change; later, after the analytical treatment had been studied, grammatical variations were introduced within the familiar utterances, such as changes of pronouns, of gender, and the like. The pressure of the memorized original patterns was so strong, and the students' analytical grasp so slight, that they were unable to master this variation technique within the small portion of drill time set aside for it.

A period of several weeks, beginning as soon as the eleven playlets had been perfected, was devoted to filming the entire series. This amateur film was intended to demonstrate some of the possibilities of similar movies which would be made by professional French actors. Since the films would show how each skit was to be staged and since their sound-tracks could serve as models for oral imitation, this dramatic technique of teaching elementary language could be used by teachers who did not have any of the rather elaborate facilities available to this experiment, the stage, the dramatics coach, and the corps of drill-masters. These and other possible values of the film are not germane to our interests here. The fact is pertinent, however, that the unprecedented amounts of time and energy which went into the process of shooting it contributed little to the students' further acquisition of the linguistic skills. Consequently this movie project clearly illustrates a point relevant to the consideration of many experimental courses: they were usually not limited to the rigid testing of the experimental procedures we are discussing here, but also incorporated various theories and interests of the local staffs, which sometimes involved work not directly related to the major purposes of the course. As a result, the present experimental programs are often far from demonstrating the maximum effect obtainable from the concentrated and efficient use of the basic procedures employed.

When, after the filming, normalcy was restored to the course, what remained of the third quarter was used for teaching the students conventional French orthography and introducing them to reading. The spelling was taught by means of regularly written texts of the playlets, so that the printed symbols could be related to familiar sequences of sound firmly implanted in the students' auditory consciousness. While no objective data are available on the effectiveness of the late transfer, observers believe that all but the poorest students mastered the essentials of regular spelling with relative ease and that few of the group found their established speech habits altered by the appearance of the written language.

In their initiation to reading the group studied the French version of Wendell Willkie's *One World (Le monde est un)*. Very little was covered before the end of the first year, but the book was continued and finished in the second year. Instead of the conventional vocabulary or dictionary, the English version itself was used for guidance and reference. The objective was direct mastery of French literary vocabulary and expression rather than translation or structural analysis. At the same time seven contact hours per week afforded the students continuous opportunity to extend the scope of their oral production and aural comprehension well beyond the limits of the original playlets which had laid the foundation of these skills. After completing the Willkie they went on to other readings, and studied a conventional grammar of written French, to complete their second year.

The group was tested by the Investigation, along with all other groups, at the end of their first year and twice in their second year. After the first-year dramatics course they were not expected to compare favorably with any other group, since their one special skill was not measurable by either of the standard tests. We report their first-year scores in order to show the progress made during the second year:

CO-OPERATIVE

	JUNE 1945 1 YR. ADV. R		MAR. 1947 1 YR. & 2 QU. ADV. O	JUNE 1947 2 YRS. ADV. Q	
	MEAN	(%ILE)	MEAN	MEAN	(%ILE)
R	29.0	(0)	53.7	70.5	(63)
V	53.4	(26)	69.3	75.6	(63)
G	46.1	(19)	55.9	67.0	(73)
T	42.2	(7)	60.5	73.2	(67)

AURAL

	JUNE 1946 1 YR. LOWER A		MAR. 1947 1 YR. & 2 QU. UPPER B		JUNE 1947 2 YRS. UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)	MEAN	(TENTH)
I	4.1	(5)	9.5	(9)	11.9	(7)
II	11.5	(5)	16.8	(8)	21.6	(8)
III	4.8	(2)	6.4	(7)	5.1	(8)
T	20.4	(4)	32.7	(9)	38.6	(8)

The Co-operative scores indicate steady progress, through the second year, from almost no reading ability whatever to a better-than-average, well-balanced control of all three areas measured by the test.

The first-year aural scores are at the median on Parts I and II, but the expected inability to comprehend anecdotes is confirmed by the Part III score. In the second year, however, aural performance is consistently high.

Both sets of second-year scores contrast so favorably with the first-year scores as to suggest that the group's highly developed skill in acting skits, although nothing measurable by the tests themselves, nevertheless constituted an excellent foundation on which a larger body of skills was rapidly and efficiently built.

Group 1c. The elementary French experimental group of 1946-1947, numbering twenty-eight, was divided evenly into two groups to be taught independently of each other by different methods. Both operated on seven-hour-per-week schedules throughout the year.

Subgroup 1c.1 took the same dramatics course as Group 1b the preceding year. In view of the regular teaching staff's previous experience with the course, the services of the dramatics coach were dispensed with, but otherwise the procedure was essentially unchanged. However, no movie was made, and thereby nearly a full quarter's time was saved in which more use was made of the analytical and supplementary materials, transition to normal orthography was effected earlier, and more reading was done.

Subgroup 1c.2 used as their basic material Harris and Lévêque's *Conversational French,* a newly published textbook developed from conversational materials first produced in connection with a semi-intensive experimental course at the beginning level.

Both subgroups were tested at the end of the second and third quarters, with the following results:

	CO-OPERATIVE				AURAL			
	MARCH 1947 2 QU. ADV. O	JUNE 1947 1 YR. ADV. Q			MARCH 1947 2 QU. LOWER B		JUNE 1947 1 YR. LOWER A	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
	Subgroup 1c.1							
R	36.1	50.2	(33)	I	6.7	(6)	6.7	(8)
V	63.6	68.6	(83)	II	16.4	(7)	16.4	(8)
G	47.4	57.4	(63)	III	6.5	(4)	14.0	(8)
T	48.9	59.6	(61)	T	29.6	(7)	37.1	(8)
	Subgroup 1c.2							
R	47.5	56.9	(59)	I	9.9	(8)	7.8	(9)
V	62.8	68.2	(82)	II	16.1	(7)	17.7	(8)
G	58.4	66.8	(91)	III	8.1	(5)	13.9	(8)
T	56.7	65.4	(81)	T	34.1	(8)	39.4	(8)

Both the Co-operative and the aural scores of Subgroup 1c.1 are much higher than those of Group 1b at the same stage, thus indicating the effect of increased efficiency, reduced waste of time and energy, and strengthened transition to reading the second year the course was given.

At the same time the scores of Subgroup 1c.2 run generally higher than those of 1c.1, especially in Reading and Grammar. Use of the *Conversational French* thus brought the students to a higher level of general one-year achievement than did the dramatics materials; but in view of the rapid general progress made by Group 1b during the second year, it is impossible to judge the relative solidness of these one-year foundations without second-year data on both subgroups.

Group 2a. This was a group of eleven freshmen entering the B.A. program in the fall of 1944 with previous high-school study of French. During the first quarter they took the general linguistics course, and in the second quarter began an intermediate experimental French course meeting eight hours per week for a total of 140 hours over and above their previous high-school study. This was the course based principally on work with the play *Les Jours Heureux*, as described above in connection with Group 1a's second-year course (1945-1946), and its first trial was with this group in 1944-1945.

In 1945-1946 Group 2a took an advanced course meeting eight hours per week in the first quarter and five hours per week thereafter for a running total of 300 college hours. The results of their testing over the two-year period are as follows:

CO-OPERATIVE

	JUNE 1945 2 YRS. ADV. O		DEC. 1945 2 YRS. & 1 QU. ADV. R	JUNE 1945 3 YRS. ADV. R	
	MEAN	(%ILE)	MEAN	MEAN	(%ILE)
R	73.3	(73)	86.4	86.8	(88)
V	78.1	(73)	78.8	81.4	(50)
G	71.5	(87)	67.5	74.0	(83)
T	76.3	(78)	82.5	83.9	(78)

AURAL

	DEC. 1945* 2 YRS. & 1 QU. UPPER B		JUNE 1946 3 YRS. UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)
I	10.0	(6)	17.5	(7)
II	17.0	(5)	29.3	(7)
III	9.0	(7)	6.7	(6)
T	36.0	(7)	53.5	(8)

*Not tested aurally in June 1945.

Group 2a shows far higher Co-operative scores than Group 1a at the same stage (see p. 181). The difference is significant at the 5% level, and doubtless means that this group's high-school training had formed a solider basis for the development of the visual skills than Group 1a was able to establish in its elementary (experimental) course.

In the aural scores, there is no satisfactory point for comparison between Groups 2a and 1a.

Groups 2b and 2c. These groups of eight and six students entered the B.A. program as freshmen in the falls of 1945 and 1946 respectively. They took the same intermediate experimental course as Group 2a, but began in the first quarter at seven hours per week (taking one hour of linguistics concurrently). Over and above their high-school study they thus accumulated 180 hours at the end of the year, slightly more than Group 2a's 140 hours.

4

We did not follow Group 2b and 2c into their third (advanced) year. Their scores for the intermediate year are as follows:

	CO-OPERATIVE				AURAL			
			Group 2b					
	DEC. 1945 1 YR. & 1 QU. ADV. R	JUNE 1946 2 YRS. ADV. R			DEC. 1945 1 YR. & 1 QU. UPPER B		JUNE 1946 2 YRS. UPPER A	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	77.4	88.2	(98)	I	6.1	(6)	13.1	(7)
V	74.8	77.1	(69)	II	16.5	(8)	20.8	(7)
G	72.5	69.8	(81)	III	6.1	(7)	5.2	(8)
T	77.6	81.1	(90)	T	28.7	(8)	39.1	(8)
			Group 2c					
	MARCH 1947 1 YR. & 2 QU. ADV. O	JUNE 1947 2 YRS. ADV. Q			MARCH 1947 1 YR. & 2 QU.		JUNE 1947 2 YRS.	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	63.7	75.8	(80)	I	10.7	(9)	14.2	(8)
V	73.2	78.8	(76)	II	20.0	(9)	16.7	(5)
G	65.7	68.7	(80)	III	9.0	(9)	6.7	(9)
T	68.7	77.0	(80)	T	39.7	(9)	37.5	(8)

These high-school trained groups join with Group 2a in significantly surpassing college-trained Group 1a at the two-year stage in the reading skills. We also find here aural comparisons with Group 1a (not available for 2a), but in this field the difference favoring the present groups is not significant.

Comparison with Conventional Groups (French)

We tested all conventional French classes at College J in 1944-1945 and 1945-1946, but none in 1946-1947. These courses met four hours per week in the first two years for totals of 105 hours per year, and three hours per week in the third year for a total of 80 hours, bringing the three-year accumulation to 290 hours. As stated above, however, it will not be pertinent to compare experimental and conventional groups at comparable levels in terms of accumulated hours. The announced purpose of the experimental courses was not to attain the same goals in the same number of hours, but to reach a comparable level of reading skill and higher levels of oral-aural skill in the same number of years, through the device of converting unsupervised outside study into directed practice with the spoken form of the language. It is therefore on equality of calendar time that we base our comparisons between the experimental and the conventional groups.

We do not give scores of conventional groups in this section; all are contained in Appendix B. The table which follows lists the four experimental groups of 1944-1945 and 1945-1946 described above, with separate columns for the successive Co-operative and aural testings. In each of these columns appears the difference in the means between the experimental group and its conventional counterpart *on the same test at the same calendar stage.* A plus sign before the difference indicates that it favors the experimental group; a minus sign indicates that it favors the conventional group. At the final testing, the symbol ^{1} or ^{5} appearing after the difference indicates that it is significant at the 1% or the 5% level as the case may be.

In reading skills, the college-trained experimental Group 1a was significantly poorer than its conventional counterpart even at the end of two years. On the other hand the intermediate experimental groups 2a and 2b, previously trained in high school, were not significantly different from the corresponding conventional groups at two years or at three years.

In aural comprehension, only one experimental group out of four was significantly better than its conventional counterpart — namely Group 2a at three years. Nevertheless with the exception of Group 1b, which is to be considered a special case, the other (non-significant) aural differences consistently favor the experimental groups.

TABLE XI

Differences between Means for Experimental and Conventional Groups — French

		CO-OPERATIVE			AURAL		
Group 1a		1 YR.	2 YRS.		1 YR.	1 YR. & 1 QU.	2 YRS.
	R	-17.8	-13.3	I		-2.6	+0.9
	V	- 4.3	- 7.1	II		+2.5	-0.3
	G	-10.6	- 6.1	III		+3.1	+0.5
	T	-11.3	$- 9.3^5$	T		+3.0	+1.1
Group 1b		1 YR.	2 YRS.		1 YR.	1 YR. & 1 QU.	2 YRS.
	R	-28.7		I	-1.7		
	V	- 8.5		II	-2.1		
	G	-12.7		III	-4.1		
	T	-18.2^1	*	T	-7.9	*	*
Group 2a		2 YRS.	3 YRS.		2 YRS.	2 YRS. & 1 QU.	3 YRS.
	R	-0.8	-4.2			+2.3	+4.3
	V	-1.3	-0.7			+1.1	+8.2
	G	+2.[illegible]	+1.9			+1.9	+1.8
	T	-0.2	-1.0		§	+5.3	$+14.3^5$
Group 2b		2 YRS.				1 YR. & 1 QU.	2 YRS.
	R	+9.6		I		+1.1	+2.7
	V	+1.8		II		+8.1	+3.2
	G	+1.1		III		+1.8	+1.6
	T	+4.4		T		+11.0	+7.5

*Conventional group not tested.

§Neither group tested.

There is not in these figures a strongly enough defined pattern to indicate decisively that the French experimental program *as a whole* did or did not achieve its announced objectives.

Spanish

Group 1a. This group of seventeen B.A. students took the elementary experimental course in 1944-1945, and nine of them continued with the intermediate course the following year. Like the corresponding French Group 1a (and for the same reason: see p. 180) they began Spanish in the second quarter and met for lecture and drill eignt hours per week for a total of 180 contact hours.[19] Their principal textbook was Krakeur's *Las Aventuras de Roberto Martín,* the Decca Language Series method with accompanying phonograph records reproducing the text. Much supplementary material in mimeographed form was also used. Lectures by the course director to the full group dealt with the linguistic structure of Spanish, with the phonograph records used for illustration of the phonology; drill work was devoted to mimicry and memorization of the text — a mystery story with background circumstances selected from the usual stock of travel, food, sport and amusement situations — and to practice with a number of supplementary exercises, based on that text, which the course director himself produced. The group was split into two sections for the drill work, which was conducted by Latin Americans attending the university in one or another capacity, who were not by profession teachers of Spanish. These drill-masters rotated from one to the other drill section every other week, in order that the students might accustom themselves to hearing more than one native speaker's voice and accent.

During the three quarters of the second year Group 1a met seven hours per week[20] and accumulated a two-year total of 390 hours. They worked principally with Tyre's *Speaking Spanish,* supplemented by mimeographed materials.

The nine students in Group 1a who completed two years of Spanish[21] produced the following test scores:

---

[19] Three hours per week were devoted to lecture, five to drill.

[20] Three in lecture and four in drill.

[21] See Appendix B for scores from the original seventeen at one year.

CO-OPERATIVE

	JUNE 1945 1 YR. ADV. O		DEC. 1945 1 YR. & 1 QU. ADV. O	JUNE 1946 2 YRS. ADV. O	
	MEAN	(%ILE)	MEAN	MEAN	(%ILE)
R	60.6	(76)	64.0	70.6	(69)
V	54.8	(47)	56.6	65.9	(46)
G	40.7	(11)	49.9	61.6	(46)
T	52.3	(41)	57.8	67.6	(55)

AURAL

	DEC. 1945* 1 YR. & 1 QU. UPPER B		JUNE 1946 2 YRS. UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)
I	13.1	(8)	15.2	(9)
II	13.0	(8)	26.1	(7)
III	4.4	(3)	5.4	(6)
T	30.5	(7)	46.7	(8)

*Not tested aurally in June 1945.

The Co-operative scores indicate that the group *(a)* was adequate in Reading at the end of one year and maintained its level in the second year; (*b*) stood slightly below the median in Vocabulary in both years; (*c*) was very weak in Grammar at one year but progressed nearly to the median in the second year.

The aural scores show (*a*) excellent comprehension of the difficult upper-level material after only one quarter of the second year, except on the Dialogue (in which respect cf. French Group 1a), and (*b*) development of high aural competence in the second year.

Group 1b. This was the 1945-1946 counterpart of Spanish Group 1a, made up of twenty-five entering B.A. freshmen. Materials and procedure in both the first and the second year were the same as for Group 1a, but on the revised basis of seven hours per week throughout both years,[22] divided into 3 of lecture and 4 of drill. The test results from this group were as follows:

---

[22] See Footnote 18, page 182.

CO-OPERATIVE

	JUNE 1946 1 YR. ADV. O		DEC. 1946 1 YR. & 1 QU. ADV. O	JUNE 1947 2 YRS. ADV. O	
	MEAN	(%ILE)	MEAN	MEAN	(%ILE)
R	59.8	(73)	64.3	68.9	(63)
V	53.7	(41)	56.7	66.9	(50)
G	56.3	(66)	63.7	67.3	(69)
T	57.4	(62)	62.8	69.6	(63)

AURAL

	JUNE 1946 1 YR. LOWER A		DEC. 1946 1 YR. & 1 QU. UPPER B		JUNE 1947 2 YRS. UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)	MEAN	(TENTH)
I	8.4	(8)	11.7	(8)	11.2	(6)
II	12.4	(6)	8.8	(6)	23.5	(6)
III	14.5	(6)	7.7	(7)	6.2	(7)
T	35.3	(7)	28.2	(7)	40.9	(7)

As contrasted with Group 1a, this group learned grammar adequately in the first year, and its Vocabulary and Grammar levels were higher at the end of two years, dropping nowhere below the median.

Group 1b's aural ability, well-developed at one year, remained consistent but was one decile lower than that of Group 1a at the end of two years.

Group 1c. The elementary Spanish experimental group of 1946-1947, numbering twenty-seven, did not differ in respect to method or time-schedule from Group 1b of the preceding year. The results of March and June testings were as follows:

	CO-OPERATIVE				AURAL			
	MARCH 1947 2 QU. ADV. O	JUNE 1947 1 YR. ADV. O			MARCH 1947 2 QU. LOWER B		JUNE 1947 1 YR. LOWER A	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	56.7	63.3	(84)	I	11.9	(9)	8.2	(8)
V	53.2	54.1	(43)	II	14.5	(8)	12.0	(6)
G	49.8	56.1	(65)	III	14.7	(5)	15.0	(6)
T	53.7	58.9	(68)	T	41.1	(8)	35.2	(7)

Group 1c, weakest in Vocabulary just as Group 1b, surpassed the latter group in Reading and Vocabulary, and matched it in Grammar, at one year. On the aural test, Group 1c's one-year scores are almost identical, part for part, with those of Group 1b. The drop from 8th to 7th tenth from March to June is probably due merely to the difference of forms and does not represent a regression in aural competence.

The observable rise in performance level on the Co-operative tests by the three successive first-year groups, without a compensatory drop in the aural scores, suggests that the experimental method was successfully perfected in the achievement of reading and grammar objectives as the teaching staff gained experience.

Group 2a. This was a group of twenty-three freshmen entering the B.A. program in the fall of 1944 with previous high-school study of Spanish. During the first quarter they took the general linguistics course, and in the second quarter began the same intermediate Spanish course already outlined as taken by Groups 1a and 1b in their second year, but meeting eight hours per week for a total of 140 hours over and above their previous high-school study.

In 1945-1946 thirteen of the original twenty-three continued Spanish by taking an advanced course meeting eight hours per week in the first quarter and five hours per week thereafter for a running total of 300 college hours. The results of testing the thirteen continuing students over the two-year period are as follows:

CO-OPERATIVE

	JUNE 1945 2 YRS. ADV. 0		DEC. 1945 2 YRS. & 1 QU. ADV. 0	JUNE 1946 3 YRS. ADV. 0	
	MEAN	(%ILE)	MEAN	MEAN	(%ILE)
R	76.2	(87)	84.0	80.3	(76)
V	76.1	(82)	78.3	82.4	(75)
G	64.1	(56)	64.6	73.8	(76)
T	74.0	(79)	76.7	81.5	(79)

AURAL

	JUNE 1945 2 YRS. UPPER A		DEC. 1945 2 YRS. & 1 QU. UPPER B		JUNE 1946 3 YRS. UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)	MEAN	(TENTH)
I	17.2	(9)	14.3	(5)	19.2	(9)
II	31.1	(9)	13.7	(5)	32.3	(9)
III	7.2	(8)	8.2	(4)	7.7	(9)
T	55.5	(9)	36.2	(5)	59.2	(9)

Group 2a shows considerably higher Co-operative scores than Groups 1a and 1b at the same stage (see pp. 196-197). This difference, although not significant, suggests as in the case of French that high-school training may have furnished a solider basis for the development of the reading skills than the college groups had been able to establish in their elementary (experimental) course.

As for the aural scores, this group was one decile higher at two years than Group 1a, and two deciles higher than Group 1b, but the differences are non-significant. The drop from the 9th to the 5th tenth of the upper-level test between June and December 1945 is difficult to explain, although it is doubtless partially attributable to the difference of forms.

Groups 2b and 2c. These groups of twenty-seven and thirty-eight students entered the B.A. program as freshmen in the falls of 1945 and 1946 respectively. They took the same intermediate experimental course as Group 2a, but began in the first quarter at seven hours per week. Over and above their high-school study they thus accumulated 180 hours at the end of the year, slightly more than Group 2a's 140 hours. Except as stated below, we did not follow Groups 2b and 2c into their third (advanced) year. Their scores for the intermediate year:

Group 2b

	CO-OPERATIVE				AURAL			
	DEC. 1945 1 YR. & 1 QU. ADV. O	JUNE 1946 2 YRS. ADV. O			DEC. 1945 1 YR. & 1 QU. UPPER B		JUNE 1946 2 YRS. UPPER A	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	68.2	72.2	(75)	I	10.8	(5)	15.2	(9)
V	67.6	73.4	(74)	II	9.8	(7)	27.7	(8)
G	61.0	70.9	(81)	III	8.2	(7)	4.1	(5)
T	67.2	74.2	(80)	T	28.8	(7)	47.0	(8)

Group 2c

	DEC. 1946 1 YR. & 1 QU. ADV. O	JUNE 1947 2 YRS. ADV. O			DEC. 1946 1 YR. & 1 QU. UPPER B		JUNE 1947 2 YRS. UPPER A	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	70.8	74.1	(80)	I	12.0	(8)	15.3	(9)
V	67.0	73.5	(75)	II	11.3	(7)	26.1	(7)
G	67.6	69.8	(78)	III	7.7	(7)	6.9	(8)
T	70.4	74.6	(80)	T	31.0	(7)	48.3	(9)

These high-school trained groups join with Group 2a in surpassing college-trained Groups 1a and 1b at the two-year stage on the Co-operative tests, but these differences and those on the aural are non-significant.

Twelve of the twenty-seven members of Group 2b continued with the advanced course in 1946-1947, and were tested aurally once during that year, but not visually. Their scores on Aural Upper B in December 1946 (at 2 yrs. and 1 qu.) were:

	MEAN	(TENTH)
I	15.3	(5)
II	15.9	(6)
III	9.1	(5)
T	40.3	(6)

Comparison with Conventional Groups (Spanish)

As in the case of French, we tested all conventional Spanish classes at College J in 1944-1945 and 1945-1946, but none in 1946-1947. The total number of hours accumulated by these Spanish classes was the same as in French (see p. 192). However, for the reasons stated above in connection with French (see

p. 193), we compare the experimental and conventional groups only on the basis of equal calendar time.

Actual scores of conventional Spanish groups are found in Appendix B. The following table lists the four experimental groups of 1944-1945 and 1945-1946 described above, with differences in the Co-operative and aural between the experimental group and its conventional counterpart on the same test at the same calendar stage.[23]

TABLE XII

Differences between Means for Experimental and Conventional Groups — Spanish

		CO-OPERATIVE			AURAL		
Group 1a		1 YR.	2 YRS.		1 YR.	1 YR. & 1 QU.	2 YRS.
	R	-2.2	-1.6	I		+5.2	+6.0
	V	-7.0	-8.0	II		+5.6	+1.8
	G	-15.5	-7.3	III		-1.4	+0.7
	T	-8.9	-6.2[5]	T	$	+9.4	+8.5
Group 1b		1 YR.	2 YRS.		1 YR.	1 YR. & 1 QU.	2 YRS.
	R	-5.0		I	+2.4		
	V	-7.4		II	+0.4		
	G	-6.3		III	+0.7		
	T	-6.8[1]	*	T	+3.5	*	*
Group 2a		2 YRS.	3 YRS.		2 YRS.	2 YRS. & 1 QU.	3 YRS.
	R	-2.6	+1.7	I		+1.4	+4.0
	V	-5.6	-2.9	II		+0.4	+3.1
	G	-0.9	-0.4	III		+1.4	+2.2
	T	-3.6	-0.7	T	*	+3.2	+9.3
Group 2b		2 YRS.	3 YRS.			1 YR. & 1 QU.	2 YRS.
	R	+0.9		I		+4.3	+5.2
	V	-0.6		II		+3.8	+4.4
	G	+2.8		III		+2.3	-0.1
	T	+1.0	*	T		+10.4	+9.5[1]

*Conventional group not tested.
$Neither group tested.

---

[23] For full explanation of the figures see that for the corresponding French table, p. 193.

In reading skills, the college-trained experimental Group 1a was significantly poorer than its conventional counterpart even at the end of two years, as was also Group 1b at the end of one year. On the other hand the intermediate experimental Groups 2a and 2b, previously trained in high school, were not significantly different from the corresponding conventional groups at two years or at three years.

In aural comprehension, only one experimental group out of four was significantly better than its conventional counterpart — namely Group 2b at two years. Nevertheless the other (non-significant) differences tend to favor the experimental groups by small margins.

The pattern in these figures is strikingly similar to that of the French comparisons between experimental and conventional groups, and like the French pattern it does not shed strong light on the success of the Spanish experimental program *as a whole.*

German

German groups were followed by the Investigation only through the year 1945-1946. The German staff worked from the beginning of experimentation to develop an integrated two-year program with a three-fold purpose, involving the attainment of: (1) a reading knowledge for the effective use of German in connection with various fields of learning, (2) reasonable competence in aural comprehension and oral production of the language, (3) some acquaintance with the background and culture of the German people. The textual material for the two-year course was prepared by the staff.

Group 1. Sixteen freshmen entering the B.A. program in the fall of 1945 were enrolled in the elementary experimental German course. Like the French and Spanish groups of that year, they met in class seven hours per week (plus one concurrent hour of linguistics) for a total of 180 hours. The lecture hours (two per week) were devoted to the presentation of fresh material and the analysis of grammatical structure; the five drill-hours were used for oral and aural practice with the material previously presented in lecture. Two hours of out-of-class study were required of the students. All staff members were native or bilingual speakers, and were experienced teachers of German. No mechanical audio-visual aids were employed.

Group 1 was tested both visually and aurally twice during the year, with the following results:

	CO-OPERATIVE				AURAL			
	MARCH 1946 2 QU. ADV. O	JUNE 1946 1 YR. ADV. Q			MARCH 1946 2 QU. LOWER A		JUNE 1946 1 YR. LOWER A	
	MEAN	MEAN	(%ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	49.5	54.2	(42)	I	6.8	(6)	9.9	(8)
V	51.1	55.5	(46)	II	5.6	(4)	7.5	(6)
G	49.6	54.1	(55)	III	8.8	(5)	11.4	(6)
T	50.3	55.0	(48)	T	21.2	(6)	28.8	(7)

In the reading skills these students were below the one-year median in all except Grammar. In aural ability they were slightly above average by the end of two quarters; their third-quarter rise can be accepted only with reservations as far as progress is concerned, since the same form of the test was repeated.

Group 2. Nine freshmen entering the B.A. program with previous high-school study of German took the intermediate experimental course. This was scheduled at seven hours a week (three of analysis and four of drill), and four hours of outside study were expected of the students. This group was tested four times: at the beginning of the first quarter and at the end of each quarter. We report scores for all testings except the third.[24]

[24] Scores for all testings are given in Appendix B.

CO-OPERATIVE

OCT. 1945		DEC. 1945	JUNE 1946	
1 YR.		1 YR. & 1 QU.	2 YRS.	
ADV. Q		ADV. Q	ADV. Q	
MEAN	(%ILE)	MEAN	MEAN	(%ILE)
51.3	(31)	59.6	65.7	(44)
50.2	(26)	58.7	65.8	(43)
52.0	(47)	59.1	62.4	(50)
51.1	(33)	59.9	66.0	(45)

AURAL

	OCT. 1945		DEC. 1945		JUNE 1946	
	1 YR.		1 YR. & 1 QU.		2 YRS.	
	LOWER A		UPPER A		UPPER A	
	MEAN	(TENTH)	MEAN	(TENTH)	MEAN	(TENTH)
I	7.4	(6)	5.6	(5)	13.7	(9)
II	7.7	(6)	20.6	(8)	29.6	(9)
III	7.6	(5)	7.0	(6)	11.4	(9)
T	22.7	(6)	33.2	(8)	54.7	(9)

Like the elementary students, this intermediate group stood below the median on all scores for the Co-operative except Grammar. In aural ability they were on a par, at entrance from high school, with college-trained Group 1 at the end of two quarters. After one quarter they had made noteworthy aural progress: 8th tenth on the upper-level test. By the end of two years they placed in the 9th tenth, but this was on the *third* testing with the same form of the same test; they had taken it a second time at the end of the second quarter (March) and had then already reached the 9th tenth for total score.

Comparison with Conventional Groups (German)

Conventional German groups were tested in 1945-1946 at identical stages with the experimental groups. The total number of hours accumulated by these German classes was the same as in the other languages (see p. 192). However, for the reasons stated above in connection with those languages (see p. 193), we compare the experimental and conventional groups only on the basis of equal calendar time.

The scores of the conventional German groups are given in Appendix B. The following table shows the differences in the mean scores between the experimental group and its conventional counterpart on the same test at the same calendar stage.[25]

TABLE XIII

Differences in Means between Experimental and Conventional Groups — German

		CO-OPERATIVE			AURAL		
Group 1		1 YR.			1 YR.		
	R	-3.1		I	+3.1		
	V	-4.7		II	+1.9		
	G	+4.4		III	+2.6		
	T	-1.1		T	+7.6		
Group 2		1 YR.*	2 YRS.		1 YR.*	1 YR. & 1 QU.	2 YRS.$
	R	-1.0	-1.3	I	+1.0	+1.1	+3.1
	V	-1.0	+1.4	II	+2.0	+7.2	+13.6
	G	+0.8	+1.8	III	-0.4	+1.7	+4.3
	T	-0.6	+0.5	T	+2.6	+10.0	+21.0[1]

*At entrance from high school

$For the experimental group, this is the March rather than the June testing.

In reading skills, the experimental groups did not differ significantly from their conventional counterparts at one year or at two years.

In aural comprehension, the intermediate experimental group significantly surpassed the corresponding conventional group at two years, and other (non-significant) differences are at least consistently in favor of the experimental group.

These figures constitute evidence that in 1945-1946 the German experimental program was successful in terms of its objectives.

---

[25] For full explanation of the figures, see that for the corresponding French table, p. 193.

Summary

College J provides our most valuable body of data for comparison of results achieved in an experimental oral-aural program with the attainments of conventionally taught students. Both types formed part of an essentially homogeneous student body; although the type of language instruction they received was determined according to the curriculum (A.B. or B.S.) they were following, the average of their intelligence, previous language background and general scholastic record differed little or none. Furthermore, while the A.B. groups took semi-intensive language courses providing oral-aural training through the utilization of increased in-class practice time, both types devoted approximately equal amounts of over-all time (including home study) to language and were thus measurable side by side, for progress and ultimate achievement, throughout the two- or three-year span of their study.

The announced objective of the experimental program was to give the students a degree of oral and aural skill surpassing that imparted in conventional courses, and while devoting less time to direct instruction in reading, to produce equally high proficiency in that skill by the end of two years as an assumed consequence of early oral-aural command. The results of our testing give little evidence that this dual objective was impressively or consistently attained. In aural comprehension, only 3 out of 9 experimental groups performed significantly better than their conventional counterparts at the end of two years. In reading (including vocabulary and grammar), two experimental groups were significantly poorer than the conventional groups at two years. Where differences were non-significant, they tended to favor experimental group in aural comprehension and conventional groups in reading. The aural differentials were the least that could be expected as a consequence of direct training. On the other hand the reading outcomes did not bear out in convincing fashion the theory that an initial oral-aural command (of the still very imperfect kind possessed by these groups) can be substituted for direct training in that skill.

College B

College B, at which we extensively tested conventional classes and reported the results in Chapter V (p. 134 ff.), also instituted a series of experimental intensive courses, beginning with the academic year 1944-1945, in French, German, Portuguese, Russian, and Spanish. The Department of Romance Languages administered the experimental courses in all the languages mentioned except German, and in collaboration with that department we tested various intensive French, Spanish and Russian groups in 1945-1946 and 1946-1947.

The intensive program was established at College B in response to what was felt to be renewed insistence, resulting from the war experience, on maximum possible attainment in the *use* of a foreign language studied. The program for each language was set up as a four-year sequence in which qualified students who so desire may obtain a departmental major. The sequence constitutes in a very real sense a "language-and-area" program, since the materials employed for learning and practicing the use of the language, whether it be in reading, writing, or speaking, consist wholly of information about the area in which that language is spoken.

Enrollment in the intensive program at College B is restricted to highly qualified students genuinely interested in acquiring second-language proficiency as implementation of their cultural studies or future careers. Only those are admitted who have either a superior previous record in foreign-language study or a very high general average, or who demonstrate by tests a high linguistic aptitude. As a consequence, intensive enrollment up to the present time has been extremely small. These circumstances make the program atypical as contrasted with the full-scale conversion to intensive instruction at College H, the fairly extensive experiment at College J, or the practice at some colleges of offering intensive courses on a purely optional basis. The results achieved in this atypical College B program might have important implications for the more generally prevalent type, were it not necessary to interpret with extreme caution the data yielded by the very small number of students involved.

Each of the first- and second-year intensive courses is under the direction of a particular staff member who gives the grammatical instruction and co-ordinates the conversational practice, given by a native-speaking instructor, with other phases of the course. The amount of instruction given by the director varies somewhat from course to course, but tends to be about three hours per week. Several of the directors are bilingual, and they begin conducting grammatical discussion in French or Spanish well before the end of the first semester of the elementary course.

Practice groups for drill and conversation normally contain no more than ten students. In the opinion of the Department fully satisfactory native speakers to conduct these sessions can be recruited only from among experienced teachers — not necessarily language teachers, but persons who understand the implications of the teacher-learner relationship and the importance of lesson material carefully prepared in advance of class.

The French and Spanish courses were identically organized and may be treated together; Russian courses will subsequently be treated separately.

French and Spanish

Groups 1a and 1b. These groups represent the beginning intensive course, open only to students with no previous experience in the language. It operated on a schedule of nine contact hours per week, thus accumulating 135 hours each semester for a total of 270 hours in the first year. The students, expected to do appropriate amounts of out-of-class study, received six hours' credit per semester, i.e., double the credit for a conventional three-hour course. Their achievement at the end of the year is therefore measurable on a normal two-year basis.

Actually these groups were tested at points nearer the beginning than the end of their second semester. This being the case, their Co-operative scores are measured against the one-year norms; but their aural scores, made as they were during an "intermediate" semester, are subject to comparison with the advanced norm-group for the lower-level test.

Group 1a is composed of the students of 1945-1946; 1b, those of 1946-1947. Both were extremely small. Their scores:

CO-OPERATIVE AURAL

French 1a (N=5)

	FEB. 1946 (150 HRS.) ADV. R			FEB. 1946 (150 HRS.) LOWER B	
	MEAN	(1 YR. %ILE)		MEAN	(TENTH)
R	52.2	(40)	I	10.6	(9)
V	58.8	(47)	II	17.8	(8)
G	57.4	(63)	III	13.6	(7)
T	56.8	(50)	T	42.0	(9)

Spanish 1a (N=8)

	ADV. Q				
R	65.1	(88)	I	13.3	(9)
V	65.3	(84)	II	15.0	(8)
G	65.0	(91)	III	20.5	(8)
T	66.6	(90)	T	48.8	(9)

French 1b (N=7)

	MARCH 1947 (180 HRS.) ADV. R			MARCH 1947 (180 HRS.) LOWER A	
	MEAN	(1 YR. %ILE)		MEAN	(TENTH)
R	50.1	(32)	I	7.9	(7)
V	62.1	(60)	II	11.1	(3)
G	59.4	(70)	III	13.9	(6)
T	57.6	(53)	T	32.9	(6)

Spanish 1b (N=12)

		ADV. Q			
R	64.4	(87)	I	13.1	(9)
V	63.5	(79)	II	16.3	(9)
G	69.1	(96)	III	15.0	(6)
T	67.3	(91)	T	44.4	(9)

On the Co-operative tests, all were strongest in Grammar, a situation often characteristic of students taking an intensive course of the current type. In French, both groups in Reading, and Group 1a in Vocabulary, had not yet passed the one-year median despite their 15 to 45 hours of post-elementary instruction.

On the aural side both Spanish groups and French 1a showed superior performance, while French 1b, with 25 more hours, failed to reach comparable levels. The discrepancy can be due merely to individual student differences in these very small groups.

Groups 2a and 2b. These were students who, having completed two, three, or four years' high-school study of the language, took an intermediate intensive course meeting six hours per week, 180 hours for the year. As in the beginning course, double credit was given. Their achievement at the end of the year is thus measurable on a three-year basis.

Group 2a, that of 1945-1946, was tested twice: the first time shortly after the beginning of the course, and the second time at 140 hours on the 180-hour scale. This group comprised only three students in French and five in Spanish. Their scores

	CO-OPERATIVE					AURAL			
	French 2a (N=3)								
	OCT. 1945 (30 HRS.) ADV. R		APR. 1946 (140 HRS.) ADV. R			OCT. 1945 (30 HRS.) UPPER A		APR. 1946 (140 HRS.) UPPER B	
	MEAN	(1 YR. %ILE)	MEAN	(3 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH
R	63.0	(79)	71.0	(35)	I	8.0	(5)	5.7	(2)
V	69.7	(85)	73.0	(18)	II	15.0	(4)	12.0	(4)
G	68.3	(93)	74.7	(85)	III	5.0	(8)	5.7	(4)
T	68.7	(89)	75.0	(44)	T	28.0	(5)	23.3	(3)

Spanish 2a (N=5)

	ADV. Q		ADV. Q					(A)	
R	73.6	(98)	79.8	(75)	I	15.8	(9)	21.4	(10)
V	73.4	(97)	89.0	(91)	II	35.6	(10)	38.6	(10)
G	76.8	(99)	79.4	(90)	III	10.2	(10)	9.2	(9)
T	77.0	(99)	85.6	(89)	T	61.6	(10)	69.2	(10)

Obviously, however, little weight can be attached to the performance of three or five students at any given time.

Group 2b, the 1946-1947 version of this same course, numbered eleven students in French and twelve in Spanish. They were tested once, at 120 hours on the 180-hour scale, with the following results:

	CO-OPERATIVE			AURAL	
	French 2b (N=11)				
	MARCH 1947 (120 HRS.) ADV. R			MARCH 1947 (120 HRS.) UPPER A	
	MEAN	(2 YR. %ILE)		MEAN	(TENTH)
R	80.7	(90)	I	16.3	(9)
V	79.2	(76)	II	28.8	(9)
G	71.1	(86)	III	4.4	(7)
T	77.2	(81)	T	49.5	(9)
	Spanish 2b (N=12)				
	ADV. Q				
R	68.4	(61)	I	11.9	(8)
V	72.4	(72)	II	26.2	(7)
G	75.4	(91)	III	5.3	(6)
T	74.4	(80)	T	43.4	(7)

This French group (which is of a more significant size) appears to have been considerably superior to French 2a, and even if measured by the three-year norms (which they need not be at only 120 hours), their scores would still surpass the others on both tests. In Spanish, on the other hand, the situation is reversed, with tiny Group 2a showing notably better results than the larger 2b in both fields.

Groups 3a and 3b. These were the students who, having completed the 270-hour beginning course, took in the second year of the intensive sequence a course which was distinct from that taken by the groups entering the program with high-school experience (2a, 2b). This course met only four hours per week (giving four semester-hours' credit), thus adding 120 hours and bringing the two-year total to 390 hours. Their achievement at the end of this time is measurable on a three-year basis. The course was given only in Spanish in 1945-1946 (Group 3a) and 1946-1947 (Group 3b).

Group 3a was tested twice: once shortly after the beginning of the course, and again at 350 hours on the 390-hour scale. There were nine students.[26] Their scores:

CO-OPERATIVE AURAL

Spanish 3a (N=9)

	OCT. 1945 (280 HRS.) ADV. Q		APR. 1946 (350 HRS.) ADV. Q			OCT. 1945 (280 HRS.) UPPER A		APR. 1946 (350 HRS.) UPPER A	
	MEAN	(2 YR. %ILE)	MEAN	(3 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	75.4	(84)	79.4	(73)	I	16.1	(8)	18.4	(9)
V	85.9	(97)	92.3	(95)	II	26.1	(6)	35.1	(9)
G	76.9	(93)	79.8	(91)	III	5.2	(6)	8.6	(9)
T	82.1	(95)	86.9	(91)	T	47.4	(7)	62.1	(10)

These figures, especially in the field of the reading skills, speak for themselves.

Group 3b comprised four of the eight members of the Spanish 1a group of the preceding year. They were tested once, also after 350 of their 390 hours. The results:

CO-OPERATIVE AURAL

*Spanish 3b (N=4)**

	MAR. 1947 (350 HRS.) ADV. Q			MAR. 1947 (350 HRS.) UPPER A	
	MEAN	(3 YR. %ILE)		MEAN	(TENTH)
R	74.0	(52)	I	11.8	(5)
V	93.0	(96)	II	27.5	(7)
G	79.0	(90)	III	4.8	(6)
T	85.0	(87)	T	44.1	(7)

*Two of the four failed to take the Co-operative Test.

---

[26] Their Ohio Psychological decile mean was 6.9 (cf. p. 141).

In comparing these students' scores on the aural test with those of Group 3a, it should be noted that they took it only once, while the others' 350-hour scores are the result of a second exposure to the same form.

Group 4. These students were taking the third course of the intensive sequence, which is designed as a literary course meeting three hours per week. It is distinct from the conventional literature courses in that a large amount of material can be thoroughly covered — entirely in the foreign language — not only because of the students' greater proficiency in handling the language itself, but also because of the greater amount of introductory background in literature and civilization gained in the preceding courses of the intensive sequence. Simultaneously with the courses in literature, the students take a two-hour-per-week course in conversation for one hour of credit.

We tested Group 4 only in 1945-1946, during which year all its members, in both French and Spanish, had formerly completed the course for students with high-school experience, in which they had accumulated 180 hours within the program. The advanced course (counting the conversation phase) added 150 hours, bringing the total to 330. The group was tested twice: once shortly after the beginning of the course, and again at 280 hours on the 330-hour scale. There were three students in French, and four in Spanish of whom two are excluded from the data as being native speakers of Spanish from Puerto Rico. The scores for Group 4,[27] which at the second testing are measured by the four-year norms, are as follows:

---

[27] Decile means on the Ohio Psychological Test: French 8.0, Spanish 6.5.

French 4 (N=3)

	CO-OPERATIVE					AURAL			
	OCT. 1945 (195 HRS.) ADV. R		APR. 1946 (280 HRS.) ADV. R			OCT. 1945 (195 HRS.) UPPER A*		APR. 1946 (280 HRS.) UPPER B	
	MEAN	(3 YR. %ILE)	MEAN	(4 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	91.3	(94)	98.3	(99)	I	21.0	(9)	10.3	(6)
V	82.0	(52)	86.7	(53)	II	35.7	(9)	20.3	(7)
G	76.7	(89)	70.0	(58)	III	8.3	(7)	8.7	(7)
T	87.0	(86)	88.7	(80)	T	65.0	(9)	39.3	(7)

Spanish 4 (N=2)

	ADV. Q		ADV. Q					(A)	
R	71.0	(40)	82.5	(63)	I	14.5	(7)	21.0	(10)
V	90.0	(92)	96.0	(93)	II	26.0	(6)	35.0	(9)
G	81.5	(94)	78.5	(79)	III	6.0	(7)	8.5	(9)
T	83.5	(85)	89.0	(85)	T	46.5	(7)	64.5	(10)

*Six months earlier these students had taken a tryout form of Aural Upper A, with the material spoken by their instructor.

Russian

In Russian, with the exception of a one-year special course of three hours in scientific Russian, all instruction is given within the intensive program.

Group 1. The beginning course in Russian is conducted, as in French and Spanish, on a basis of nine contact hours per week for a total of 270 hours in the first year. Credit is for six semester-hours.

The beginning Russian group was not tested in 1945-1946. The twenty members of the 1946-1947 beginning course were tested at 180 hours, with the following results:

Russian 1 (N=20)

	CO-OPERATIVE			AURAL	
	MARCH 1947 (180 HRS.) INVESTIGATION'S TEST			MARCH 1947 (180 HRS.) LOWER A	
	MEAN	(TENTH)		MEAN	(TENTH)
R	18.0	(6)	I	9.3	(6)
V	12.7	(5)	II	13.1	(5)
G	16.0	(5)	III	14.0	(4)
T	46.7	(5)	T	36.4	(5)

Groups 3a and 3b. These were students who, after having completed the 270-hour beginning course, took in the second year a course meeting six hours per week (for six credits),[28] thus adding 180 hours and bringing the two-year accumulation up to 450 hours.

Group 3a (1945-1946) was tested twice with the aural test only: once shortly after the beginning of the course, and a second time at 400 hours on the 450-hour scale.[29]

AURAL
Russian 3a (N=6)

	OCTOBER 1945 (290 HRS.) LOWER A MEAN	(TENTH)	APRIL 1946 (400 HRS.) UPPER A MEAN	(TENTH)
I	13.2	(7)	10.7	(5)
II	16.7	(7)	22.0	(5)
III	24.5	(8)	10.5	(7)
T	54.4	(8)	43.2	(6)

Group 3b (1946-1947) was tested once at 400 of the 450 hours, as was Group 3a.

Russian 3b (N=4)

CO-OPERATIVE	MARCH 1947 (400 HRS.) INVESTIGATION'S TEST MEAN	(TENTH)	AURAL	MARCH 1947 (400 HRS.) UPPER A MEAN	(TENTH)
R	19.5	(6)	I	8.8	(4)
V	17.5	(8)	II	18.5	(4)
G	15.5	(5)	III	1.5	(1)
T	52.5	(6)	T	28.8	(2)

Summary

By and large the results attained in the intensive program at College B, as measured by the Investigation's tests, are seen to be superior in both the reading and the aural field.

---

[28] As contrasted with only four hours (and four credits) in French and Spanish.

[29] Their Ohio Psychological decile mean was 6.0.

The students had been carefully selected for qualifications and interest, and they constitute our best example of what can be accomplished on a limited basis which rules out students of mediocre aptitude or slight motivation. Unfortunately for our studies, the numbers of qualified and interested students receiving intensive instruction at College B were so small that the results must be subject to cautious interpretation.

College C

As stated in Chapter V, we tested language groups at College C in both 1945-1946 and 1946-1947. We reported the results from the first of those years in Chapter V because the courses were all conventional. The data from the second year are treated here because at that time one new experimental section was instituted both in French and in Spanish.

The experimental course was intended to bring students without previous experience in the language up to a two-year level of competence in one year of study. No mere telescoping of two conventional years into one, it was further designed to test the assumption that, by making speech the initial approach, it would be possible in the same amount of time to develop aural and oral skills more than is normally done in conventional courses and yet to attain the same degree of reading ability.

Of the nine contact hours per week, six were devoted to classroom work with the full group meeting together — for explanation of grammatical structure and for some oral exercise — under the course director who was himself a native speaker. Three more hours a week were used for conversation sessions led by a native speaker who was not a trained teacher. For these conversation hours the full group of about fifteen students was divided into separate sections of five each. These three conversation hours supplanted home study and required no outside preparation. For the six class hours, which did require extra study, six credit hours were given and in that light the course was regarded as essentially the double of a conventional course. Thus the 135 hours of the first semester are to be equated with the 90 hours of one regular year, and the full 270 hours with the 180 hours of two regular years.

In theory, the intensive course was open to students who were above average in aptitude and had had no previous experience in the language. In actuality, a few of the students were average or less in aptitude, hardly any were brilliant, and several were veterans who had studied the language slightly some months or years earlier.

	CO-OPERATIVE			AURAL	
	French (N=13)				
	JUNE 1947 (270 HRS.) ADV. P			JUNE 1947 (270 HRS.) UPPER B	
	MEAN	(2 YR. %ILE)		MEAN	(TENTH)
R	75.7	(80)	I	12.2	(10)
V	73.5	(57)	II	16.5	(8)
G	66.5	(69)	III	6.2	(7)
T	74.2	(70)	T	34.9	(10)
	Spanish (N=15)				
	ADV. O				
R	70.6	(70)	I	14.1	(9)
V	67.7	(54)	II	17.1	(10)
G	72.1	(84)	III	8.2	(7)
T	72.2	(72)	T	39.4	(10)

The Spanish section had previously taken the Lower Level Form A aural test along with conventional groups in January, 1947, at 135 hours. The mean of fourteen of the fifteen students on the total for that test was 41.8, which places them in the 9th tenth of the one-year norms.

At the 270-hour (two-year) level, the French section is seen to have been strongest in Reading, the Spanish section in Grammar, with both weakest in Vocabulary. The over-all performance in the two languages is strikingly similar, and while good in the reading skills is definitely superior in the aural field.

Comparison with Conventional Groups

Since conventional groups were also tested in June, 1947, we have valuable comparative data against which to measure the attainments of the experimental group. The scores made by the regular groups appear in Appendix B. The following table sets forth the differences between the means of the intensive group and those of the regular second-year group which was at the same time completing 180 hours of conventional instruction. A plus sign means that the difference favors the experimental group; a minus sign that it favors the conventional.

TABLE XIV

Differences between Means of Experimental and Conventional Groups — College C

	CO-OPERATIVE		AURAL	
French	R	+ 6.1	I	+ 6.2
	V	- 0.5	II	+ 3.7
	G	+ 0.8	III	+ 1.4
	T	+ 2.1	T	+11.3[1]
Spanish	R	+ 2.7	I	+ 3.4
	V	+ 1.1	II	+ 8.1
	G	+ 6.2	III	+ 1.5
	T	+ 3.7	T	+13.0[1]

On the January aural testing at 135 hours, the Spanish experimental group surpassed its conventional (90-hour) counterpart by 14.0 on the total score, a difference which is significant at the 1% level.

It is thus objectively demonstrated that the goals of the experimental course were successfully attained. Moreover the similarity of the differences for both languages suggests that they are not the result of chance, but rather of the effectiveness of teaching procedures which are shown to justify the theory on which they were based.

College K

Parallel to the conventional elementary courses, experimental intensive courses were instituted at College K (a large state university) by the departments of Romance Languages and German in the year 1944-1945. In co-operation with these departments, the Investigation followed experimental and conventional groups in French, Spanish, and German during 1945-1946, the second year of the intensive program's operation. Inasmuch as the intensive courses did not follow an identical pattern for all languages, we present the data for each separately.

One principle common to all the languages, however, was that the intensive courses were not designed to telescope the coverage of two years into one, but rather to convert the student's normal out-of-class study into supervised drill in the presence of a native speaker. The objective of the experimental course was therefore the acquisition of the usual reading ability plus a better developed control of the spoken language than is normally imparted in conventional instruction. Regular courses met five hours per week, and five credits were assigned for either regular or intensive courses. Students were free to elect either type of course.

French

The elementary group. The experimental elementary course in French, as organized in 1945-1946, provided fifteen contact hours per week: five class hours of regular instruction in grammar and reading (duplicating in this respect the conventional course), plus ten hours of conversation practice. Eighteen out of the sixty-eight students enrolled in beginning French elected the intensive section. They were tested at the end of the first quarter when they had accumulated 180 hours, and at the end of the second quarter at 360 hours. Their scores on the two testings were as follows:

French (N=18)

	CO-OPERATIVE				AURAL			
	DEC. 1945 (180 HRS.) ADV. R	MARCH 1946 (360 HRS.) ADV. R			DEC. 1945 (180 HRS.) LOWER A		MARCH 1946 (360 HRS.) UPPER A	
	MEAN	MEAN	(1 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	41.1	65.1	(84)	I	2.4	(3)	4.8	(3)
V	55.2	67.6	(80)	II	8.4	(3)	15.9	(5)
G	53.4	60.2	(73)	III	8.9	(4)	2.6	(6)
T	49.9	65.6	(82)	T	19.8	(4)	23.3	(3)

In March, having devoted 120 of their 360 hours to class instruction in reading, the group comfortably reached the top quarter of the one-year Co-operative norms. In the aural field, however, their performance was surprisingly low in view of the hours devoted to oral-aural practice. Their performance on the upper-level aural test is an interesting exemplification of the inability of first-year students to cope with difficult material despite many hours spent in listening to a native speaker. Their relatively high score on Part III (Dialogue), however, suggests that their general comprehension may have been ahead of their vocabulary.

The conventional counterpart of this group was tested at the end of the year, at 180 hours. Their total means reached the 93rd percentile of the Co-operative and the fifth decile of the lower-level aural norms. Actually we should have tested the experimental group at the end of the year also, instead of in March. Because of this inequality in time, there is no real comparability; but it seems not unlikely that in one more quarter the experimental group could have made up the difference on the Co-operative test and would have surpassed the others on the *lower-level* aural by a fairly wide margin.

Spanish

Group 1. As contrasted with French, elementary Spanish provided only ten contact hours per week: the five of regular instruction, plus five of conversational drill. Twenty-six out of the total first-year enrollment of ninety-four chose the intensive group, which was tested at the same calendar points as the French, namely at the end of the first quarter at 120 hours, and at the end of the second quarter at 240 hours. Their scores on the two testings:

CO-OPERATIVE AURAL

Spanish 1 (N=26)

	DEC. 1945 (120 HRS.) ADV. Q	MARCH 1946 (240 HRS.) ADV. Q			DEC. 1945 (120 HRS.) LOWER A		MARCH 1946 (240 HRS.) UPPER A	
	MEAN	MEAN	(1 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	45.5	52.3	(43)	I	3.7	(4)	6.6	(4)
V	40.5	46.7	(19)	II	8.0	(3)	18.4	(4)
G	42.4	55.3	(61)	III	14.5	(6)	4.1	(5)
T	42.2	51.7	(40)	T	26.2	(4)	29.2	(4)

On the Co-operative tests, comparison with the corresponding French group does not favor the Spanish, as the number of hours devoted to reading instruction was the same for both groups (120 hours by March) and the Spanish passed the median only in Grammar. On the aural side, however, the Spanish group reached almost exactly the same levels as the French (at least in terms of our deciles), both in December and in March, with only one-half the time devoted to oral-aural drill.

The conventional Spanish elementary group was tested at the end of the year, at 180 hours. Their total means stood at the 90th percentile of the Co-operative and the fifth decile of the lower level aural norms. It seems doubtful whether in one more quarter the intensive group would have made up the difference on the Co-operative test, but they would have probably surpassed the others easily on the lower-level aural test.

Group 2. These eleven students had completed one year of intensive elementary instruction in 1944-1945. When tested in December, 1945, they were finishing the first quarter of an intensive intermediate course operating on the same schedule of ten hours per week distributed evenly between regular class instruction and conversational drill. They had accumulated at this point a running total of 480 hours. Their scores:

Spanish 2 (N=11)

	CO-OPERATIVE			AURAL	
	DEC. 1945 (480 HRS.) ADV. Q			DEC. 1945 (480 HRS.) UPPER A	
	MEAN	(2 YR. %ILE)		MEAN	(TENTH)
R	67.6	(58)	I	13.5	(9)
V	66.4	(48)	II	26.8	(7)
G	68.7	(76)	III	3.5	(5)
T	67.5	(56)	T	43.8	(8)

This group, with 240 hours of instruction in reading during one year and one quarter, had already passed the two-year median in all scores except Vocabulary. Their aural performance is good, though not exceptional considering the number of practice hours accumulated.

German

The elementary group. Intensive instruction in German was given only at the first-year level. After completing the elementary course, students continuing in German entered the regular intermediate course unsegregated from those who had begun with the conventional instruction. As in Spanish, the intensive elementary course in 1945-1946 provided for ten contact hours per week: five of regular class instruction, each followed by an hour of conversational drill under a different (native) instructor.[30] Fourteen out of sixty-three students enrolled in beginning German elected the intensive course. Both the regular and the experimental groups were tested simultaneously at the end of the second quarter, and the latter was tested again at the end of the year. We list the scores of both groups, because in this case we shall be able to make close comparisons.

---

[30] Both regular and intensive classes read Fröschel's *Himmel, meine Schuhe!* (Heath); in addition, while the regular classes used Burkhard's *Lernen Sie Deutsch!* (Holt), the intensives worked with Rehder and Twaddell's *Conversational German* (Holt), a text especially adapted to oral-aural procedures.

German Experimental Group (N=14)

	CO-OPERATIVE					AURAL			
	MARCH 1946 (240 HRS.) ADV. Q		JUNE 1946 (360 HRS.) ADV. Q			MARCH 1946 (240 HRS.) LOWER A		JUNE 1946 (360 HRS.) UPPER A	
	MEAN	(1 YR. %ILE)	MEAN	(1 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	51.1	(31)	56.1	(49)	I	7.4	(6)	6.6	(7)
V	47.5	(18)	57.0	(53)	II	6.6	(5)	18.1	(7)
G	52.6	(51)	56.9	(67)	III	8.6	(5)	7.9	(7)
T	50.3	(29)	57.1	(57)	T	22.6	(6)	32.6	(8)

German Conventional Group (N=49)

	MARCH 1946 (120 HRS.)			MARCH 1946 (120 HRS.)	
R	52.9	(38)	I	4.8	(5)
V	50.5	(28)	II	4.4	(3)
G	55.0	(59)	III	7.8	(5)
T	53.0	(40)	T	17.0	(4)

In March although the conventional group was ahead of the experimental by small margins, the difference in total-scores proved non-significant. At the same stage the experimental group bested the conventional on Parts I and II of the aural test; but on Part III (Anecdotes) they were unable to excel, presumably because of their sparse vocabulary (cf. their low Co-operative Vocabulary score). The difference between the aural totals was also non-significant.

Substantially all the students in both German groups had taken, on entrance to College K, the ACE College Aptitude Test of 1937. The raw-score means of the two groups on this test were

Experimental Group: 94.2
Conventional Group: 88.9

This difference too was non-significant.

Summary

At College K all the experimental groups except the Spanish elementary produced results indicative of success in achieving the reading objective. Moreover, in view of the satisfactory performance of the intermediate Spanish group also tested, it seems reasonable to infer that those students had reached an acceptable level of reading skill at the end of their first year, and that the failure of the 1945-1946 elementary group to do so may have been due to the quality of the students or to other factors rather than to the methods of instruction.

College L

A joint committee representing the Departments of Romance Languages, German, and Slavonic Languages at College L (a large state university) recommended revision of basic-level language courses in such a way as to embody certain essential features of current experimental instruction: (1) concentration of time devoted to language study, (2) reduction in the number of students per class section, (3) increased emphasis on use of the spoken language on the assumption that oral-aural proficiency is not only a valid aim in itself but that it can also facilitate the acquisition of an appreciative and durable reading knowledge. We are concerned here with changes made, as a result of the committee's recommendations, in the basic French and Spanish courses, some sections of which we tested in 1946 and 1947.

The Year 1945-1946

The first year of experimentation (1945-1946) saw established the experimental framework within which a more controlled series of operations were carried out in the following year. The traditional arrangement, which was continued for students already embarked on the fulfilment of their three-year language requirement, consisted of an elementary-intermediate sequence followed by a third-year course, all on a three-hour-per-week, three-credit basis. In 1945-1946 the initial two-year sequence was paralleled by a new five-hour, four-credit course of one year, designed to prepare students for the third-year course in five-sixths as much contact time as the earlier program and in one-half as much calendar time. We tested French and Spanish groups of both the experimental and the traditional type, but on different bases in the respective languages.

The French experimental group. The several sections of this first experimental French course spent an average of six introductory hours on pronunciation, using phonetic symbols and the exercises in Swanson's *Concise French Grammar;* thereafter they covered all the lessons of the above grammar and of Hendrix and Meiden's *Beginning French,* read *Maupassant: Six Contes Choisis* and (outside of class) five of the first six booklets of the Heath-Chicago Rapid Reading Series. Chief innovations of the course were considered to have been greater insistence on pronunciation, which was repeatedly checked and twice recorded for each student, and the use of phonograph recordings of the pronunciation drills in the grammar and of all the readings in *Beginning French.*

Both the experimental and the traditional group were given the Co-operative Test. The results were not available to us, but it has been reported that comparison favored the experimental group. Our aural comprehension test was also administered to both groups at an equal number of contact hours (110), the following scores resulting:

French	*Experimental Group (N=184)*		*Traditional Group (N=53)*	
	LOWER A		LOWER A	
	MEAN	(TENTH)	MEAN	(TENTH)
I	4.5	(5)	4.2	(5)
II	12.2	(6)	13.0	(6)
III	10.3	(5)	10.6	(6)
T	27.0	(6)	27.8	(6)

Differences are minimal on all parts of the test.

The Spanish experimental group. In the first semester this group used LaGrone's *Conversational Spanish for Beginners,* and read about forty-five pages of Roessler and Remy's *First Spanish Reader*. In LaGrone the translation exercises were usually written by the students; the questions on the reading selections were used in various ways at the instructors' discretion; the readings themselves were normally translated, and also formed the basis of conversation in Spanish. Some instructors made recordings of students' speech as an aid in teaching pronunciation. In the second semester they covered alternating assignments in Grismer and Arjona's *Short Spanish Review Grammar* and Ashburn's *Selected Spanish Short Stories.* The latter book, of which they read about seventy pages, proved too difficult, so that selected lists of words used were mimeographed to serve as a basis for active vocabulary building. Answers to the questions based on the stories were written out; there was considerable oral translation in class from Spanish to English, and frequent dictations were given.

The aural test was administered to one section of the experimental course and eight of the traditional course at the same calendar time in January, 1946. The experimental group had completed 110 hours, the traditional group about 130 hours. The scores:

Spanish	*Experimental Group (N=15)*		*Traditional Group (N=133)*	
	LOWER A		LOWER A	
	MEAN	(TENTH)	MEAN	(TENTH)
I	8.0	(8)	5.5	(6)
II	12.4	(6)	10.1	(5)
III	16.1	(7)	11.7	(5)
T	36.5	(8)	27.3	(5)

The appreciable difference between the scores of the two groups is significant.

The Year 1946-1947

During the first year of experimentation it became increasingly clear to the staff that, with or without objective data on the current operation, many questions of technique were still going unanswered. For example, what is the value of dictation? Is oral-aural drill *the* gateway to reading ability? Can a near-perfect accent be imparted to first-year students? Can pronunciation best be taught through imitation or through phonetic description? Can rote drill be used to teach grammar? What degree of grammatical accuracy should be insisted upon? In the hope of arriving at some satisfactory answers to these and many other similar questions, it was agreed by the Romance Language Department to organize and carry out experimentation in techniques within the frame of the newly established five-hour basic course. The outcomes of such experimentation were to be evaluated by different types of objective test. The University administration agreed to bear extra expenses connected with the experiments.

For the year 1946-1947, it was decided to use two approaches, systematically differentiated on the following basis: Plan A, an essentially traditional technique, using a text in normal orthography from the first day, teaching pronunciation mainly by imitation but without the supplementary use of recordings, teaching grammar inductively and with written exercises, and not avoiding translation of reading material into English; Plan B, a more radical method, not allowing the students to see the written form of the language until several weeks after the beginning of the course, insisting on accurate pronunciation taught by phonetic description and supplemented by extensive use of recordings, teaching grammar deductively as points arise in connection with reading and requiring no composition by the students, and insisting on all feasible classroom use of the foreign language from the outset.

Approximately twelve hundred students, comprising sixty sections of twenty each, took the basic course in 1946-1947. Half of the sections were taught in accordance with Plan A and the other half followed Plan B. Students were not allowed to shift from one plan to the other at mid-year.

Two hundred eighty-nine French and five hundred fifty Spanish students following Plans A and B took the Co-operative and aural tests at 110 contact hours in March, 1947, with the following results:

CO-OPERATIVE			AURAL		
		French Plan A (N=149)			
	ELEM. R			LOWER A	
	MEAN	(1 YR. %ILE)		MEAN	(TENTH)
R	56.7	(57)	I	4.3	(5)
V	59.7	(52)	II	10.5	(4)
G	57.6	(65)	III	10.5	(5)
T	58.7	(59)	T	25.3	(5)
		French Plan B (N=140)			
R	52.7	(43)	I	4.9	(6)
V	56.7	(39)	II	10.0	(4)
G	52.4	(40)	III	9.6	(5)
T	54.3	(39)	T	24.5	(5)
		Spanish Plan A (N=287)			
	ELEM. O OR P				
R	59.0	(71)	I	5.9	(6)
V	52.9	(39)	II	10.7	(6)
G	55.3	(61)	III	14.9	(6)
T	56.4	(56)	T	31.5	(7)

Spanish Plan B (N=263)

R	60.2	(74)	I	6.7	(7)
V	53.7	(43)	II	11.2	(6)
G	53.7	(57)	III	15.3	(6)
T	56.6	(60)	T	33.2	(7)

The most striking difference is that between French Plan A and Plan B on the Co-operative test, favoring Plan A by a fairly wide margin on each of the three parts. On the total score, the 4.4 difference in means proved to be significant at the 1% level. At 110 hours out of the 150 provided in a course aiming to reach a two-year level at the end, French Plan B had not yet reached the one-year median in any part of the Co-operative test.

In Spanish the differences on the Co-operative are all non-significant.

Apart from the obvious fact that the Spanish groups give evidence of generally higher aural competence than the French, there is nothing to be observed regarding the comparative aural scores.

These are, of course, meager objective data by which to answer the question of which Plan yields better results in what kinds of foreign-language competence. One indication, seemingly valid for French but not supported by the parallel data in Spanish, points to Plan A, the more traditional technique, as superior for the teaching of reading ability, passive vocabulary control, and grammatical analysis. At least there is no evidence in this body of data to the contrary.

In any case the language staff at College L are convinced that they are on the right track and that if they are able to carry out their three years of experimentation, they will possess valid objective answers to some of the important problems of foreign-language teaching. Teachers everywhere may look forward with interest to the full reports which the staff promise to make public on their procedures and findings at the end of each year.

College M

In 1945-1946 the Department of German initiated an experiment at College M (a privately endowed college for women) in the teaching of its introductory course. Two considerations influenced the decision to try the experiment: (1) the success of the methods used in Armed Service language courses seemed to warrant their adaptation to college teaching wherever possible; (2) the traditional attempt to develop speaking, understanding, reading and writing ability simultaneously, without distinguishing between the relative importance of these skills, was increasingly felt to be producing neither a high level of achievement in any one of them nor (generally speaking) an appreciation for the living language.

As designed to fulfill the college language requirement and to serve as a prerequisite for advanced language and literature studies, the experimental course was conceived to have the following aims:

1. A reading knowledge effectively usable in later German courses and in other fields.
2. A fair degree of facility in speaking and comprehending German.
3. An understanding of the structure of the language.
4. Some practice in expression in written German.

A report from the course director goes on to state:

> Since most students who study German in college do so in order to acquire a reading knowledge, this aim had to be of major importance and the course had to be planned with this in mind. It seemed probable, however, that this objective could be accomplished more effectively if emphasis were put from the beginning on that most natural approach to any language, the spoken word. Besides, most students, particularly women, have always had a lively interest in learning to speak and understand a foreign language, even when their

ultimate aim has been the acquisition of a reading knowledge. In order to speak and comprehend and to read fairly complicated texts it is essential that the student learn to see how the language "works," that is, to understand what is generally included under "grammar," and, beyond that, to acquire some understanding of word formation, sentence structure and meaning, and, if it is not too presumptuous to hope to accomplish in an introductory course, to gain some feeling for style and some insight, no matter how slight, into the *Geist* of the language.

We decided, therefore, to approach the learning of German through speaking and understanding the language. The development of these skills was to dominate the work of the first semester. Reading was to be introduced as early as possible and was to take up an increasing amount of time in the second semester. Analysis of the structure of the language was to accompany and form a part of the work of both semesters. Writing was to occupy a relatively small amount of the student's time throughout the year.

In order to insure a registration large enough to make the experiment possible, Freshmen and Sophomores were excluded from the conventional three-hour elementary course, which continued open to Juniors and Seniors only. Forty-four students were enrolled in the experimental course.

Organization was on a basis of nine contact hours per week: three of structural analysis and six of drill. For the analysis classes the group was divided into two sections of twenty-two each; and for drill there were several sections of not over nine students each, most of which were conducted by one drill instructor at the beginning of the week and by another at the end.

Each student was expected to spend a minimum of nine hours a week in out-of-class study. During the first semester extra evening drill sessions were arranged for those students who desired to do their preparation under staff supervision. In the second semester a special section, replacing three hours of drill, was arranged for successful students who wished to cover more than the regularly assigned readings.

The staff of the course consisted of the director (an American) and four drill instructors, two of whom were native Germans. The director taught the sections in grammatical analysis, conducted weekly staff meetings, prepared assignment schedules, corrected all the written work done by the students, and determined grades after consultation with the drill instructors.

The basic text, used throughout the year, was Rehder and Twaddell's *Conversational German*. In the first semester the method of mimicry-memorization was used exclusively in learning the units of the text. Presentation of a given unit was made in analysis class, and that unit was drilled in the practice sessions of the same week, to the point where the phrases and sentences could be reproduced naturally and unhesitatingly and where, after a few weeks, variations on the memorized patterns could be successfully produced. The first three units were learned entirely through hearing and imitating the instructor, the students not being permitted to see any written German until the third week. No attempt was made, however, to use any system of phonetic transcription. For the first seven units the German text was read only after the unit had been committed to memory and the students, well prepared in advance for the sight of the printed text, had little trouble in mastering its spelling peculiarities.

Reading of extremely simple texts was begun in the sixth week of the course, occupying two of the drill hours through the first semester and three in the second semester. The way in which the reading problem was attacked is explained in the director's report in a paragraph which merits quotation in full:

> Until the student was able to read well independently, the drill instructor first read the text aloud in German. The students repeated after the instructor either chorally or individually or both. An explanation of the meaning of words which were not under-

stood by the student was given by the instructor, in German whenever possible. When the text had been read and re-read in this way and the meaning was clear to the students, simple questions were asked and answered on the text in German. Insofar as possible, conversation and reading were combined. As the difficulty of the reading material increased, it became evident that the active vocabulary and the patterns of speech acquired in the learning of the "mim-mem" units were inadequate for answering questions on the texts read. *Pending further study of more effective methods of applying the conversational method to the discussion of the reading material,*[31] a temporary compromise was made. In the reading of texts of advanced difficulty we decided to aim at developing as large a passive vocabulary as possible and to attempt to awaken some feeling for the literary qualities of the text. Some of the passive vocabulary passed naturally into the active vocabulary of the students, but in each case the amount was a wholly individual matter, depending not only on the frequency of occurrence of the word or construction but also on the receptivity of the student. Study of word formation which is so important in German was begun early in the year and continued throughout. Meanings of words, differences of usage, etc. were also emphasized with constant reference to the phrases in which they had occurred. Throughout there was as little as possible translation from the German into English.

Throughout the course, nearly five hundred pages of text were read, from the Heath-Chicago German Series through Kästner's *Die Verschwundene Miniatur* to Puckett's *Intermediate German Readings*.[32]

The writing of German was introduced gradually after three weeks in connection with the basic textbook. By the second semester the students were writing on an average of once a week — not translating English to German by grammatical rule, but attempting to render certain English passages of the textbook in the German patterns already drilled and varied in class. In the latter part of the course these written assignments were short *Aufsätze* on some phase of the stories read and discussed in class. Students were warned against attempting to write anything which was beyond their experience, were discouraged from using a dictionary, and were judged primarily on their ability to express themselves in clear, simple, and fairly natural German.

---

[31] Italics ours.

[32] The special section referred to above substituted Röseler's *Deutsche Novellen des Neunzehnten Jahrhunderts* for the last-mentioned text, and covered over a hundred pages more than the regular students.

The aural test was given to the group in March after 200 hours, and both the Co-operative and the aural in May at the end of the 270-hour course. Thirty-nine of the original forty-four students took the full battery of tests. Their performance at the end of the year is measured, naturally, against norms for two years of study. The scores:

	CO-OPERATIVE		AURAL				
	MAY (270 HRS.) ADV. Q			MARCH (200 HRS.) LOWER A		MAY (270 HRS.) UPPER A	
	MEAN	(2 YR. %ILE)		MEAN	(TENTH)	MEAN	(TENTH)
R	71.4	(66)	I	14.1	(9)	9.5	(8)
V	69.2	(57)	II	13.4	(8)	24.2	(9)
G	78.3	(94)	III	14.6	(7)	7.6	(7)
T	75.3	(79)	T	42.2	(9)	41.2	(9)

The group placed in the top quarter of the two-year norms on the strength of their score in Grammar. Despite the large amount of reading covered in the course and the attempt made by the staff to develop the students' passive vocabulary, (see p. 233), they did not reach so high a level of competence in these sections of the test as in Grammar, which was studied thoroughly under the course director in connection with the basic material of the *Conversational German* used from the beginning to the end of the course.

The effectiveness of the aural-oral approach as handled in the course is directly reflected in the aural scores, which are superior on both levels of the test.

None of the group had done any previous study of German in secondary school. All but two had studied either another modern foreign language or Latin for two to four years, and all but three of the thirty-nine had taken a C.E.E.B. Language Achievement Test for entrance to the college.[33] Their mean scores on those tests were as follows:

French - 25 cases, average study 3 yrs. : 592.2
Spanish - 3 cases, average study 2-2/3 yrs: 587.3
Latin - 8 cases, average study 3-1/4 yrs: 557.0

All these scores are seen to be above the C.E.E.B. median of 500.

The coefficients of various correlations of the Investigation test scores with each other and with three items of supplementary data are shown in the following table:

TABLE XV

College M: Correlations

		CO-OPERATIVE Q	SAT VERBAL	SECOND SEMESTER GRADE	GRADE-POINT AVERAGE
German Experimental Group (N=39)	SAT VERBAL			.46	.42
	CO-OPERATIVE Q		.52	.76	
	AURAL UPPER A	.63	.49	.62	

Partial Correlations:

1) Co-operative Q and Aural Upper A, with SAT-score held constant: .50
2) Co-operative Q and Second Semester Grade, with SAT-score held constant: .69
3) Aural Upper A and Second Semester Grade, with SAT-score held constant: .51
4) Aural Upper A and Second Semester Grade, with Co-op. Q held constant: .28

The correlation between Co-operative and aural is high. Even with the intelligence factor, as measured by the SAT-score (the College Board's Scholastic Aptitude Test) held constant, it is .50, which is the highest of these particular partial correlations to be found anywhere in our data. This group was more than usually successful on both the Co-operative and

---

[33] One was an advanced-standing transfer student from another college.

(especially) the aural tests, and the high correlation figure (in which intelligence is not a strong factor) would suggest that the success in both fields was due to effective teaching of both types of skill side by side. The same assumption is borne out by the remaining correlation figures. While course-grades seem to have rewarded the visual skills (doubtless including writing) somewhat more than the aural (-oral), the Aural-with-Semester Grade correlation does not drop sharply from the Co-operative-with-Grade correlation (cf., for example, Colleges A and B, pp. 131, 142); and moreover when the Co-operative score is partialed out, the coefficient (.28) is not strikingly low.

This pattern, as we interpret it, does not lend weight to the assumption made in some quarters that aural skills automatically show high correlation with visual skills, even if they are taught as secondary aims or not stressed at all. To us what is reflected here, rather, is the successful teaching of both types of skill as parallel objectives; and, in reflecting just that, they are in contrast with the lower correlation figures we have found in programs where either (*a*) aural development is a secondary aim of conventional instruction (cf. College B) or (*b*) both oral-aural and reading ability are leading aims of experimental instruction (cf. College H) in which, however, a more effective method of teaching reading has yet to be worked out.

College N

A small-scale experiment in elementary French was undertaken at College N (a large state university) in 1945-1946 and was repeated in 1946-1947 on a partially revised basis. The conventional elementary course was retained both years, and enrollment in the experimental course was limited, though not on a selective basis, being merely the result of chance arrangement of class sections.

The experiment was not designed to bring students of French to a two-year level in one year, but rather to produce, in terms of one year objectives, a degree of oral-aural competence higher than that normally achieved in the regular program, and an equal reading ability. The traditional first-year course provided five class hours per week, for which ten hours of outside study were expected. The experimental course called for eight class hours weekly and only four of home study, thus demanding (theoretically) less of the students' total time than the regular course. The eight contact hours were distributed evenly between four of class and four of drill in small sections, with no outside preparation required for the latter. The drills were conducted by native speakers of French who were regular members of the departmental teaching staff.

Group 1a. The 1945-1946 experimental group, twenty-five in number, used as its basic text for oral-aural work Denoeu and Hall's *Spoken French,* following prescribed procedure,[34] and for the transition to reading employed DeSauzé's *Commençons à lire*. The conventional beginning group used Pargment's *Initiation à la langue française*.

The experimental group was given the Investigation's aural test in April, 1946, at 200 hours, and the Co-operative Test at the end of the year, at 240 hours. Ninety-three students in the regular course were also given both the Co-operative and the aural at the end of the year, at 150 hours. We list the scores of both groups:

---

[34] For an outline of this procedure see College H (p. 164) and College J (p. 180).

	CO-OPERATIVE			AURAL	
French Experimental Group 1a (N=25)					
	MAY 1946 (240 HRS.) ADV. R			APRIL 1946 (200 HRS.) LOWER A	
	MEAN	(1 YR. %ILE)		MEAN	(TENTH)
R	57.7	(62)	I	5.7	(7)
V	63.6	(68)	II	11.1	(5)
G	58.2	(65)	III	8.0	(4)
T	60.8	(67)	T	24.8	(5)
French Regular Group (N=93)					
	MAY 1946 (150 HRS.)			MAY 1946 (150 HRS.)	
R	63.7	(81)	I	3.2	(4)
V	63.5	(68)	II	10.6	(5)
G	63.3	(82)	III	5.4	(3)
T	64.9	(80)	T	19.1	(3)

While the performance of the experimental group is well above the median in all parts of the Co-operative test and might be regarded as representing an acceptable level of attainment, the scores of the regular group are significantly superior (at the 5% level) in Reading and Grammar as well as on the total.

On the other hand, on the aural test, the total mean difference of 5.7 favoring the experimental group is likewise significant at the 5% level.[35]

The coefficients of correlation between the Co-operative and aural totals were .78 for the experimental group, and .47 for the regular group. The first figure is strikingly high, but the number of cases is very small.

The foregoing data indicate that the 1945-1946 experiment in elementary French at College N succeeded in turning out students who could understand (and possibly also speak) French significantly better than conventionally taught students, but in so doing failed to develop their ability to read as well.

Group 1b. In 1945-1946 the experimental elementary course in French was repeated, with the substitution of Harris and

---

[35] We report, for what it is worth, that the experimental group was given the same form (Lower A) of the aural test a second time at the end of the year and made a total mean score of 31.6 (8th tenth).

Lévèque's *Conversational French* for the *Spoken French* of Denoeu and Hall used the preceding year. Meanwhile the conventional course was divided into two groups using as texts, respectively, Micks and Longi's *Fundamentals of French* and Lecompte and Sundeen's *Unified French Course*. The experimental group of fifteen students, as well as the full conventional group, were given both the Co-operative and the aural tests at the end of the year, each group at that time having accumulated the same number of contact hours as in the previous year. Unfortunately the Co-operative scores had not yet been put at our disposal when we went to print, and we can only report the aural scores, which were as follows:

AURAL
MAY 1946 (240 HRS.)
LOWER A

	MEAN	(TENTH)
French Experimental Group 1b (N=15)		
I	4.5	(5)
II	13.2	(6)
III	8.7	(5)
T	26.4	(6)
MAY 1946 (150 HRS.)		
Group Using Micks & Longi (N=219)		
I	2.8	(4)
II	9.2	(4)
III	8.2	(4)
T	20.2	(4)
Group Using Lecompte & Sundeen (N=112)		
I	2.9	(4)
II	11.2	(5)
III	8.6	(5)
T	22.7	(5)

Experimental Group 1b slightly surpassed Group 1a on two parts and on the total score, but this superiority could be the result of testing at 240 instead of at 200 hours. In any case Group 1b also maintained a slight aural advantage over the conventional groups.

EXPERIMENTS AT THE HIGH-SCHOOL LEVEL

What few experiments we know of at the secondary-school level in the development of ability to speak and understand a foreign language have not yielded concrete evidence on the effectiveness of the newer methods. In the normal high-school curriculum no provision can practicably be made for concentration of time devoted to language study; the standard American teacher cannot suddenly be supplanted by a native speaker; indeed, within the inflexible high-school system any pressure to depart from traditional procedure is met with understandable resistance from administrators hard pressed to defend even the meager place at present conceded to foreign language study. Experiments have been, therefore, extremely limited in scope or duration.

We report briefly on three such experiments, in the course of which our aural comprehension tests were administered to the pupils involved.

High School B

In February, 1946, an experiment in the oral teaching of foreign language was undertaken in one of the high schools of a large city system. For reasons it is not necessary to enumerate here, the experiment was organized at the last minute, and lack of time prevented the formulation of strictly scientific procedures and the arranging of a clearly objective situation. Factors which minimized its scientific validity were the impossibility of grouping the pupils on the basis of I.Q. or prognostic tests; the lack of suitable texts and audio-visual aids, the absence of scientific control groups, and the non-existence of a detailed course of study.

The languages chosen were French, German, and Spanish, in each of which there happened to be two class sections of beginners. In each case the section comprising the better pupils (on the basis of their general records) was designated as the experimental group, and the other as a control group.[36] All six sections met during the same 37-minute period, five days a week. Three well-equipped, progressive and co-operative young teachers were selected from within the system, and under the circumstances the most that could be done in the way of planning the courses was to rely upon the resourcefulness of these individual teachers for the conception and practice of the actual techniques of oral-aural instruction. (The German teacher was a native speaker, the others were not.)

Each of the experimental classes was supplied with a set of the Kany elementary conversation manuals. In addition, the French group used *Colette et ses frères,* the German group Betz and Price's *First German Book,* and the Spanish group Wilkins' *Primeros Pasos*. The Kany manuals served as a model for simulated conversations, a model for copying, and a source of material for memorization. The readers were used (on an average of ten minutes in each class period) to cover the grammatical points required of first-term students of the language, and as a basis for conversation on the stories and anecdotes read. There was considerable oral activity carried on in the classroom on a spontaneous basis, developed by the teacher as the occasion demanded. Much of this consisted, especially in the beginning, of questions and answers. Later, pupils were called upon to interrogate classmates, to give orders, to paraphrase, to dramatize, and to give brief descriptions or narrations. This sort of activity meant constant demands on the resourcefulness and energy of the teachers, and it is reported that they acquitted themselves well.

Emphasis was placed on accuracy in pronunciation, although there was no resort to phonetic transcriptions; rather a method of practical phonetics was employed, with frequent use of the blackboard for illustrative purposes. English was theoretically banned from the classroom. Grammatical concepts were taught by means of gesture, circumlocution, analogy, comparison, and the like; frequent and continued oral practice was employed to enforce the new patterns.

---

[36] The experimental sections contained forty-one pupils in French, thirty-six in German, thirty in Spanish. Control sections were of approximately the same size.

Oral and written tests were given at intervals to ascertain the progress of the students, and at the close of the term in June, all six classes were given the customary standard examination used in the system. On this examination the experimental classes in French and German are reported to have performed considerably better than the control groups, while the special class in Spanish failed to achieve a higher rating.[37]

The experimental groups in the three languages were maintained throughout their second term the following year (fall of 1946), and at the end of that second term ten pupils[38] from each group were given the Investigation's aural comprehension test with the following results:

	French		*German*		*Spanish*	
		LOWER A		LOWER A		LOWER A
	MEAN	(TENTH)	MEAN	(TENTH)	MEAN	(TENTH)
I	4.1	(6)	10.6	(9)	8.7	(10)
II	9.7	(6)	7.1	(8)	12.7	(10)
III	6.2	(5)	14.8	(9)	13.8	(8)
T	20.0	(6)	32.5	(9)	35.2	(10)

These scores are measured against the secondary-school norms for two terms of study. The German and Spanish results are strikingly good in terms of our norms, but unfortunately the corresponding control groups (if indeed they were kept intact at all during the second term) were not tested and we therefore lack data for comparing these levels of aural achievement with those of conventionally taught pupils in the same school.[39]

---

[37] Much of the wording of the foregoing account of experimental instruction at High School B is taken from a report prepared for an educational commission by the director of foreign languages in the city system.

[38] We do not know whether these ten constituted a random selection, or were the ten best.

High School C

In the 1945 Summer Session of a large city high-school an experiment was tried in the intensive teaching of French and Spanish at the beginning level. The purpose was to test the assumption that in a short period of highly concentrated study, which affords the necessary continuity of association with the language, pupils can be given a basic oral and aural command of the language in addition to a grammatical and reading knowledge equal to that gained by regular language pupils in the same amount of time spread over a whole school year.

Studying no other subject during the summer session, the pupils met in language class four hours a day, five days a week, for seven weeks. Their total contact time was thus equal to that of pupils having four class hours a week over a school year of thirty-five weeks.

The French and Spanish groups were under the full direction of their respective teachers and did not follow identical methods.

The French group had twenty pupils and one teacher (not a native speaker). She endeavored to use, and to have the pupils use, the spoken language as much as possible in class, but otherwise she relied rather heavily on the conventional procedures with which she had become familiar in the preparation of pupils for state language examinations. She used a conven-

[39] In another school of the same city system, certain groups in the same three languages were tested aurally in April, 1947. It had been planned to provide during that year double periods for the instruction of third-year classes in one of the schools, but we cannot learn whether the groups tested in April were those particular third-year classes. The total means in French and German were within a very few points of those listed above; in Spanish the total mean was 10 points lower.

tional textbook, taught grammar inductively from it, practiced little if any mimicry-memorization, and used no audio-visual aids.

The Spanish group had twenty-five pupils and two teachers — neither of them a native speaker. The one in charge had acquired some first-hand familiarity with methods and techniques employed in the ASTP and in later experiments along the same lines at the college level, and he planned the course in such a way as to embody the main features of the oral-aural approach. He selected as a textbook LaGrone's *Conversational Spanish for Beginners*, and to accompany this text he secured a set of phonographic transcriptions — recorded by a bilingual American professor of Spanish — of all the material in the book. For the introduction of new material in LaGrone, the class met as a unit and listened first to the recorded spoken version, imitating and drilling it chorally under the direction of both teachers working as a team. Only after they had become thoroughly familiar with the sound of the new material and had mastered its meaning in association with the sound, were they allowed to see the printed version. Grammar was taught in accordance with the functional treatment given it in the textbook; but the verb paradigms had been recorded along with the other material, and a goodly amount of time was spent (ill-advisedly, we feel) in the "mim-mem" of such utterances as "*Tuve, tuviste, tuvo, tuvimos, tuvisteis, tuvieron.*" After the presentation, initial mimicking and explanation of new material, the group was divided into two sections of twelve pupils or so each for further drill and conversation practice under each teacher separately.

Both the French and the Spanish group took the Investigation's aural test (lower level) at the end of their seven-week, 140-hour course. The results:

	French (N=20)		*Spanish (N=25)*	
	LOWER A		LOWER A	
	MEAN	(TENTH)	MEAN	(TENTH)
I	6.6	(9)	7.4	(8)
II	8.6	(5)	10.0	(8)
III	6.6	(6)	15.1	(9)
T	21.8	(7)	32.6	(10)

The scores are measured against the secondary-school norms for two terms of study. The results from the Spanish group are seen to be excellent in terms of those norms; and while there is no proof of it, the possibility cannot be overlooked that the higher level reached by the Spanish than by the French group directly reflects the amount of aural proficiency teachable through each of the rather sharply contrasting approaches used.

No specific information was available to us on the success of these experimentally taught students in their second year of conventional high-school language study.

High School D

In a small high school an experiment was carried out during one year (1945-1946) with a group of first-year French students. The purpose was to test the effectiveness of one of the currently available courses in which conversation is taught through the use of phonograph records. The group of seventeen pupils was divided in half, both taught by the same instructor. During the first semester, Group A worked with the phonographic materials while Group B was taught by a conventional grammar-translation method. The two groups met the same number of hours per week. During the second semester the methods were transposed: Group A moved over to grammar-translation, and Group B used the records. Both groups took our lower-level aural test at the end of each semester, with the following mean scores for the total test:

	FIRST SEMESTER LOWER A		SECOND SEMESTER LOWER B
	MEAN	(TENTH)	MEAN
Group A	21.1	(7)	20.2
Group B	14.5	(4)	13.4

While the difference between the means at the end of the first semester did not prove to be significant, that existing between the means on the second test is significant at the 1% level. No data were available to us on the intelligence or general linguistic aptitude of the pupils forming the two groups, and in the absence of such data we cannot advance any theory to explain the significant difference in the end results that would not leave several variables out of account. If the difference had been significant at the end of the first semester, before the methods were shifted — or better yet at the end of both semesters, — some support might then be found for the hypothesis that an initial approach through the spoken language serves to develop a higher aural facility (shown here to be retainable even after a shift of emphasis to visual skills, cf. Group A) than can be attained in a situation in which the emphasis is first visual and only later oral-aural (Group B).

Summary

Since the work of these experimental groups was a point of cardinal interest to the Investigation, the summary of their procedures and results forms a major part of our general conclusions in Chapter VIII.

Chapter VII

The Advanced Level of Language Skill: English as a Foreign Language

The Investigation was particularly fortunate in being able to study language learning at an advanced level. Usually such study is not possible. In American institutions few students continue beyond the second or third year of foreign language work; and for those who do, advanced work tends to emphasize literature, historical linguistics, and other matters rather than the linguistic skills. Our work at the advanced level was, consequently, not with the usual "foreign languages" but with English taught as a foreign language to the many foreign students who have come to the USA during the past few years. The Investigation profited from the opportunity to observe this work, and we believe that all language teachers can benefit from considering it. Though a different medium is involved, many of the basic problems of language teaching at the advanced level are the same. When these issues are stated in terms of one's native language, many of them become clearer and are less likely to be decided on the basis of habit or tradition.

In America the teaching of English as a foreign language had long been neglected and even scorned. Though valuable work had been done for immigrants and other groups, it was usually carried on under adverse conditions by teachers who were undertrained, underpaid, and overworked. This situation began to change markedly even before the second world war.

American advances in various technologies brought increased numbers of foreigners to work and study in American universities, hospitals, social agencies, and governmental bureaus. With the coming of the war, the closing of the European universities and the inauguration of training programs for the armed forces of our allies brought still more. The postwar influx has been equally large.

These students did not labor under the handicaps of advanced age, low I.Q., or inferior educational, social, and economic status as did many earlier students of English. Still more important, instead of being isolated in the slums of our cities, this new group was thrust upon the attention of institutions and agencies which had resources in trained personnel and other facilities far beyond any heretofore devoted to English teaching. Skilled language teachers and linguists either found themselves confronted by this new problem on their own campuses or were called in by various agencies to help. As a result, there has been a degree of interest and a variety of experiment approaching that called forth in the foreign-language field by the Army Specialized Training Program.

Many of these students come with little or no English; but probably the majority of them have completed two to four years of English in high schools or colleges in their native countries. They thus represent approximately the next linguistic level above that covered in the preceding chapters; and their problems are those which our American student will meet, if after completing the two or three-year course, he has the opportunity of living and studying in the country whose language he has learned.

Work at this level has a note of realism which is impossible at the earlier stages. What skills and how much of them a student should have at the end of 100 or 200 class-hours of college French must be a somewhat arbitrary judgment. In the case of these students, however, the amount of language they need is obviously what they must have in order to carry on their work here. This pragmatic test offers some basis for standards, but it is complicated by a number of factors (many of which are equally operative in the case of American students of foreign languages at this level).

The first is that these students vary enormously in the amount of language they need to get along and in the relative command of each of the basic skills which they must possess. This variation is a function of the different uses which these students have for language as a result of the vocational fields in which they work, the types of activity involved, the conditions under which they live, and the like. For example, the foreign student who lives in a community of his compatriots and who works in a laboratory, running experiments and recording the results in numerical or chemical symbols, may get along with surprisingly little English. On the other hand, the student who must deal with American clients and patients, who must cover long lists of assigned readings in English, or who must follow university lectures, will need a rather extensive knowledge of the language.

As the preceding examples imply, there will also be some variation in the amount of each skill the student needs. Some will need great ability in reading and writing but will have relatively little call to speak or understand. Conversely, others will need primarily aural-oral skills and little else. Obviously then, no single standard of achievement can be set for all the group, and one of the first problems is to define needs of various general classes of student. The county agricultural agent or the foreman on the assembly line may have a very different standard, but in its way, a no less difficult one than that of the English professor. With certain exceptions (which will be indicated), however, we have based our findings on those who will be "students".

A second important variable is the amount of allowance which Americans will make for the foreigner in his attempt to cope with English. Some agencies, institutions, supervisors, and instructors will make great concessions. Others will adopt a a sink-or-swim policy. In smaller organizations more personal allowance is possible, and in some places standards are not so high. There the foreign student needs much less English. For one reason or another, however, foreign students tend to affiliate with the larger agencies and institutions—exactly the places where the least individualization of treatment is possible, where the demands are greatest, and where the native students are most talented. If one attempts to equate a specific level of performance by certain students with judgments as to whether or not these students' English is satisfactory, one finds the greatest disparity of judgment regarding similar cases—and sometimes even in regard to the identical student.[1] Frequently it is clear that the difference lay in whether the judge asked himself, "Does this person do all right for a foreigner?" or "How does he compare with the rest of my students?"

This point leads to the third, that many factors other than purely linguistic ones are involved in the selection, handling, and evaluation of foreign students in this country. We should not be so foolish as to maintain that linguistic considerations should be the sole criterion, or even the major one, in selecting students to come to this country or in determining their programs while they are here. On the other hand, we believe the other point of view is fallacious (even though it seems frequently assumed): that the achievement of our present students, whose selection and training is largely determined

[1] E.g., two students who had the highest test scores recorded at one training center during about a year were later reported by a large institution with very high standards as possessing inadequate English. Other students, from this same center and working in the same field, were thought by other universities to possess sufficient English though their knowledge was far inferior to that of the other two.

by political, social, and other factors here and abroad, should be taken as representing what is linguistically feasible, desirable, or inevitable. We may well wish to consider language as only one of a number of factors; but if we do so, we should not deceive ourselves as to what we are doing. Theoretically, it would be pleasant if we could settle the linguistic questions free from all other considerations. That situation is not likely to be available. Therefore we can only remind ourselves, in attempting to determine standards, that our present students, the length and type of training they receive, the conditions under which they have it, and a number of other matters are determined on nonlinguistic grounds.

Another question involved in setting standards for this "ability to get along" is whether "intelligibility" may be accepted as a substitute for "correctness." The following paragraphs taken from a theme submitted by a foreign student illustrate the point:

> Many year ago, when I was a young boy, it came to my mind the idea to visit some day the United States. At this time, I couldn't expect that I really should come to this country. The idea came probably to my mind through the influence of the movies where people are informed about America. Meanwhile I was poor and I had many handicaps to fight. Fortunately I increased my education and in 1936 I got my Degree in Law.
>
> When I was studying at the School of Law I felt the necessity to learn English in order to read some books about Constitutional Law. American and English books, besides French, are the best in this subject. I conceive the idea to improve the preliminary knowledges I had had at the high school.
>
> Then I began to read with the help of a good dictionary. At the beginning I had too much troubles, specially when I found some descriptions of the American way of life. But I was deeply interested in the subject and as much as I was reading or going deeply in my readings. I was liking both the English language and the American way of life, specially the American ideas about Democracy.

At this time, my wishes to know the United States had increased so much that I couldn't complain with the idea of never to come to this country.

Almost every line has at least one example of "bad English," not by the more pedantic rules of good usage, but in the fundamental structure and vocabulary of the language. On the other hand, it is doubtful whether any native speaker of English who would be in a position to receive such a communication would have the slightest doubt as to what was meant. In its context, even "complain" is clear. Shall we say that this student has committed scores of serious errors within a few lines and that his English is poor? Or shall we say that he has communicated his meaning well enough? As a matter of fact there is no doubt that the complete theme would receive very diverse grades at different American colleges, or from different professors in the same school.

The chief difficulty with taking intelligibility as a norm is that it will be a very fluctuating one. For example, the American well disposed toward the foreigner, familiar to some extent with the latter's native language, and fairly ingenious at guessing meanings, will understand much more than the person who is convinced, even before the foreigner begins, that he will not be able to understand him anyway because all foreigners talk and write gibberish.

Another of these complicating factors is the interlinguistic nature of many skills. These abilities tend to operate to about the same extent in all languages with which an individual works. The student who cannot remember a lecture which he has heard in his own tongue is unlikely to do better when he hears it in a foreign one. The person who writes his native language with no capitalization and in incomplete sentences or who is unable to organize a paragraph is not likely to learn these niceties while he is also struggling to master English. We acquired, rather incidentally, considerable evidence that in these respects foreign students are often far from competent in their native languages. Yet can we hold students to standards in English which they cannot attain in their native language; and can any training center hope to give all this and English too? These factors make it difficult to determine exactly what level of English the foreign student should be expected to master and these complications must be kept in mind when one examines the following data.

To ask foreigners to compete on exactly even terms with natives is to expect too much. But obviously, if the foreigner is to profit from his stay in this country, the linguistic gap between him and the native must be as slight as possible. Had a series of basic tests for all skills been in general use with native students, these would certainly have been the ones we used for testing the foreign students. Both groups could then have been measured by a common scale, and the abilities of various foreign groups could have been directly compared with existing norms for native speakers. Unfortunately, no such series of tests was available, and even the single instruments in common use in American colleges did not lend themselves to this purpose.

The available reading tests suffered from several drawbacks. Those which called for only a literal understanding of a passage were too infantile in content; the more advanced ones were too heavily weighted by items demanding inference and other inter-linguistic skills and gave too slight attention to that literal comprehension in which we were most interested. Sections of poetry, for which the foreign student generally could not be held responsible, made other tests unsuitable. In the field of aural comprehension we could find nothing. This neglect furnishes interesting commentary on our higher education. Reading tests regularly form a part of the examinations given entering students by most colleges on the sound theory that the student who has difficulty with reading will be unable to handle the assigned lists, bibliographies, and tests. Yet it is assumed that all students are prepared to cope with the many hours of lectures and discussions to which they listen. Remedial reading is a familiar phrase on every campus, but remedial listening has yet to appear. Audiometer tests are given, but they are equivalent to the use of an eye-chart as a reading test. Writing and grammar tests, in their turn, either called for niceties which should not be demanded of the foreigner or else were loaded with dubious items of correct "usage." For oral production (as opposed to oral composition or speech delivery) naturally nothing had been needed for native speakers.

In view of these difficulties, no alternative appeared to constructing our own set of tests.[2] Apart from the delay and labor involved in construction, this procedure necessitated gaining further data regarding the performance of native students on these tests. These data would plainly be difficult to secure; but under the circumstances, no other course seemed possible.

We began our search for standards and our construction of tests with the assumption that most of the existing standards tended to be too low. We came to this conclusion from having noticed the difficulties of foreign students at our own university, students who had been cleared by various training centers but who were, none the less, proving deficient in English. Our emphasis on higher standards here and in later pages should not be interpreted as disparagement of the training centers. When a person is on a long journey, it is quite possible to take him a considerable distance and still not convey him to his destination. The standards of many centers were, moreover, determined by the type of student they had, the length of time they were given for this training, and other factors over which they had no control and which, as we have already pointed out, were non-linguistic. Certainly there have been too many situations in which the length of training has been determined only by the time available between the date of the student's arrival in this country and the opening of his school or job—a period sometimes measurable in hours!

---

[2] They were developed by Miss Yolanda Leite with the assistance of the other members of the staff.

This situation can be illustrated at once if we turn to the first of the Investigation's tests, the *Elementary Structure Test*. This was prepared by Miss Hazel Mitchell during the time she was working at the central office as representative of the English Language Institute in Mexico City. The test was originally prepared to cover the first year's work in English as it was then given at the Institute and was intended for use there. When the test was available, however, several other institutions asked to use it, and we ended by taking it over for general distribution.

This test differed from our others in that it was pitched at a very elementary level. Within an extremely limited vocabulary it covers the syntax of the most elementary English: the formation of plurals, use of pronouns, sentence order in statements and questions, and the like. These are the elements of English as taught in a first-year course in a non-English speaking country. Yet this test was accepted as representing the level at which many students were when they entered centers for further study. As will be seen from the following table, the students who have taken this test. most of them shortly after arrival in this country, generally do fairly well. But we must not lose sight of its elementary nature. Though comparisons are difficult, if not impossible, it probably represents a level of difficulty somewhat below that of the grammar sections of the Advanced Forms of the Cooperative Tests, which we used for the foreign languages.

And one who is familiar with these tests will doubt whether a high score on them indicates that the student would succeed in working in France, Germany, or a Spanish-speaking country. In fact, some very rough approximations we worked out in connection with our own group at the University of Chicago seem to indicate that a score of about 135 (or 90%) on this test indicates—not that the student has a command of English—but that he has a good grounding and is in a position to profit from further instruction and practice in this country.

ELEMENTARY STRUCTURE TEST

(151 items)

Decile	*Scores*
9th	145
8th	138
7th	131
6th	125
5th	118
4th	107
3rd	99
2nd	84
1st	69
Mean	115.84

Based on the scores of 205 students in five training centers.

Reading is an important skill for these visitors. Fortunately it is one in which most of them are well prepared when they come. Reading is the skill in which they have had the most practice and which suffers least from being taught by a teacher whose acquaintance with the foreign language is slight.

Several desiderata controlled the construction of the reading test. One was that the materials should approximate the level of college reading. When the passages had been selected, they were reworked by five members of the staff, four of whom were experienced teachers of English to foreigners. Though we wanted to keep the material in its original state, a few changes were made in features that appeared unfair to foreigners; some rare words, unusual idioms, and uncommon sentence patterns were replaced by their more frequent equivalents. On the other hand, some rather uncommon words and structures (the meanings of which were derivable from the general context of the passage) were retained to give a note of realism to the reading. None the less, the materials have been slightly simplified below the level at which these students are asked to read for their work.

It also seemed desirable that the passages should be of fair length, approximately 250 words. When passages are of this size, few can appear in a single test. Hence the student is less certain to find among the small number of passages one which appeals to his personal interest and background; none the less, the realism obtained by forcing the student to work through several paragraphs of continuous material seemed to offset these disadvantages.

Another desirable feature seemed to be to make the students work with some speed. Many people can read a foreign language if they are given enough time to dig out each word and phrase in the dictionary and grammar. This procedure is completely impossible for the foreign student who must cover reading lists, case records, and the like. Though many of them are willing to devote additional hours to make up for their inevitable handicap, eventually the twenty-four hour day comes to an end. The foreigner is likely to bog down unless he can read with some ease.

Selection of topics for the passages presented the usual problem. Many studies[3] have shown that the reading skill of native speakers varies in relation to the kind of material read. If one believes that "English for automobile mechanics specializing in carburetors" is possible, then a student can justly be expected to read only in his field of specialization. We did not believe that language could be so easily compartmentalized, though this is a point to which we must return later. Even if it could be, the foreigner in this country will probably not find all his reading centered narrowly around his specialization. Consequently, the reading passages contain the type of material likely to be found in non-technical articles on current affairs and American life, and were drawn from two books and two magazine articles.

---

[3] E.g., F. P. Robinson and P. Hall, "Studies in Higher Level Reading Skills," *Journal of Educational Psychology*, XXXII (1941) 241-52 and A. S. Artley, "The Appraisal of Reading Comprehension," *ibid.*, XXXIV (1943) 55-60.

READING COMPREHENSION

(41 items)

Decile	*Scores*
9th	31
8th	26
7th	21
6th	17
5th	15
4th	11
3rd	8
2nd	4
1st	1
Mean	15.44

Based on scores of 250 students in five training centers.

The level of achievement is not high. The average student scores considerably less than half of the total number of points possible. But since the examination is not simple and since the student works without the dictionary and other aids which he would usually have, these results do not demonstrate the need for higher standards as clearly as do some which follow.

Aural comprehension is also an important skill for the foreign student; but, in marked contrast to reading, it is the one in which he usually is weakest. This is the form of English with which he has had least contact, and certainly the sounds which issued from a teacher who spoke English with a marked foreign accent bear little resemblance to those which he hears uttered by native speakers. And those Americans talk so fast and swallow half their words!

The desirability of having a formal test of aural comprehension can hardly be overemphasized. We have often heard teachers say that they could tell more by three minutes of interview than by fifteen minutes of test. There is considerable room for self-deception in this regard. As anyone knows who has used a foreign language, a "yes" and a smile at appropriate intervals, or a noncommital comment can often give the impression of satisfactory comprehension when the actuality has been far otherwise. Experience we have had with students whose comprehension has been certified on the basis of interview has made us extremely dubious of this procedure.

The selection of passages for this test involved much the same problems as those for the reading examination. We considered it extremely important that the passages offered should be long—in this case four to five minutes each. It is not enough for the student to be able to distinguish the individual phonemes, words, or sentences.[4] In fact, even the ability to understand two or three consecutive sentences will not carry him far. As any one of us knows who has ever tried to understand a foreign language, it is easy to follow utterances of a few words. One's auditory memory can easily retain that amount of material, and it can be mulled over if not understood as heard. Continuous discourse is much more difficult. The hearer feels he is being inundated by continual waves of sound. If he tries to stop and puzzle out the first sentence or two, he completely misses the next four or five.

---

[4] The difference in performance on the Phonetic Accuracy section or the short definitions and completions, as compared with the longer anecdotes and dialogues, has demonstrated this point in the foreign language tests. Cf. *supra*, pp. 89ff.

The native speaker, particularly the college professor, is usually not so thoughtful as to speak in phrases (or even groups of sentences) with long pauses in between. As the correlations on the foreign language tests show, students can understand single utterances without being able to cope with the full flood of speech as it pours on them. We have no quarrel with aural comprehension tests which involve simple, short materials. For placement, for practice, and for progress testing, tests of this type rather than ours should certainly be used. But we believe they are grotesquely out of place as final measures of the ability to comprehend the spoken language in anything but the most Pickwickian sense.

The importance of this skill for the foreign student has already been emphasized by our colleague, Hugh R. Walpole:

> The foreign student must read well, write understandably, and comprehend the English that he hears. He can, if need be, take longer than his fellow-students over his reading and writing. Listening time, however, is strictly limited; for he cannot at will turn the professor back and listen to him again. In aural comprehension he must be up to native standards.[5]

The scores made by the foreign students tested were as follows:

AURAL COMPREHENSION
(44 items)

Deciles	*Scores*
9th	27
8th	24
7th	21
6th	19
5th	16
4th	13
3rd	10
2nd	6
1st	1
Mean	15.56

Based on scores of 112 students in five training centers.

[5] H. R. Walpole, "How Much English Must a Foreign Student Know?" *The School Review*, LV (1947), 231.

Consideration of these scores, which reveal that only a third of the students get even half the items right, naturally raises the question, "How well do native students do on these tests?" This same question comes to mind in regard to all the tests, but this is the only one in regard to which a valid answer is likely to be obtainable. The other tests are just too simple for American college students to take them seriously. In grammar tests, for example, the American student is accustomed to being harassed with problems like "The man (who? whom?) I think will be elected is popular" and "The number of possible answers (is? are?) small." He is used to puzzling over these problems and disgracing himself by his choices. But confronted by "I arrived there (on? at? in?) Saturday," he thinks that the instructor is crazy or is pulling his leg. The native speaker simply has no question about items like this. If he makes a mistake, it is usually due to clerical error caused by carelessness or perhaps it results from the kind of confusion which afflicted the centipede who was asked which leg he moved first. The reading test is only a little better in this respect. The native college student is used to reading tests which demand some penetration. But when he sees before him in the test "The traders pressed forward following . . . the Mississippi, the other great rivers . . . ." he can complete the statement "The rivers——were the routes by which the traders penetrated the country." Few American students who have reading difficulty at this literal level manage to enter college. As a result, only the test of aural comprehension offers enough novelty and difficulty for us to have any hope that the native student will take it seriously.

This attitude of the student (whether native or foreign) toward the test is, of course, of fundamental importance. Tests are delicate instruments and are no better than the situation in which they are used. Some of our friends were very co-operative and administered the aural tests to groups

5

of native students. In most cases, however, they could not produce an actual testing situation. The Americans were not taking the test as they did their other examinations — i.e., as an important factor in determining their grade for a semester's or a year's work. The test is too simple to be even a kind of parlor game, the sort of interest which makes it possible to use otherwise unmotivated subjects in psychological experiments. The results obtained from such situations are hardly comparable with those to be expected under normal test conditions. For example, one native American made a score of 8, but he answered only the first eight questions and then apparently lost interest, trying no more items.[6]

Only one native group took the test under approximately normal test conditions, and these were students in an "awkward-squad," a section of college freshmen who were having language difficulty and had been placed in a remedial course. It would seem that the level of achievement attained by a group of this kind would give a fairly realistic picture of the kind of standard we must visualize for foreign students. Unfortunately the group was small, and certainly this evidence should not be regarded as conclusive. None the less, a comparison of its results with those of the foreign students is not without interest:

Decile	*Native Students*	*Foreign Students*
9th	40	27
8th	38	24
7th	36	21
6th	35	19
5th	34	16
4th	34	13
3rd	32	10
2nd	30	6
1st	28	1

---

[6] It may be that some of the scores reported here for foreign students are also too low for similar reasons, though we have no cause to think that such was the case for any students included in these norms.

The contrast between the two distributions (in which only the 9th decile for the foreigners is up to the first decile of the natives) requires no comment.

The writing skill was measured by a double test, *Writing and Structure*. In the first part, the writing test, the student was required to manipulate very simple materials actively. He was given a brief anecdote to read and was then asked to write, without access to the text, the answers to ten simple questions on the story. In the second section of the writing test he was asked to describe members of his family and to tell how he had spent the preceding Sunday.[7]

The structure part was in multiple-choice form. In ninety-five items, the student was required to select the correct word or phrase out of the three offered. These items formed part of a continuous narrative, the description by a foreign student of his arrival in the United States. We believed this would be a topic with which every foreign student would be familiar. Objection is sometimes made to tests of this type on the grounds that, for the foreigner, they may degenerate into puzzle-solving; instead of expressing ideas in his own way, he has to fit the proper bit of English into a hole in the context, judging the fit by sense and syntax. We sought to avoid that difficulty by making the correct answers the most normal and commonplace English. Aside from agreement in number and tense of the verbs, the items do not depend on this kind of maneuvering.

As was to be expected, the correlation between the scores on the two parts was relatively high.[8] It might have seemed hardly necessary to administer both types. Studies of the two types with the compositions of native students[9] have shown that the objective test is an equally valid and certainly a much more convenient measure. But several features of the foreigner's situation suggested that the cases were not exactly parallel and that both sorts of test should be tried.

---

[7] An average paper on this test is reproduced a few pages later.

[8] On a sample of 190 papers from seven institutions the coefficient was .70. The distribution was, however, obviously curvilinear and the correlation ratios were .73 and .86. The latter of these (structure on writing) was greater than the coefficient by an amount significant at the 1% level.

In general, the structure test is highly reliable and a better predictor of performance in the free writing test than the latter is of it. The explanation of these facts is fairly clear. In writing (particularly in regard to the simple material required in this test), the student can often do fairly well by carefully keeping within whatever small amount of English he knows; by using baby sentences, he can stay out of complications. This device is a handy one for anyone learning a strange language and should certainly not be penalized. One wonders, however, whether the student is equally careful when he is not in a test situation. A more dubious practice (but one in which every language learner tends to indulge) is just not to try to say what he isn't sure of being able to say correctly. The tests written on this principle are extremely laconic. Although this may be a means of doing well on tests of free composition, retribution strikes when the student later follows this plan in preparing papers in other subjects; the instructor has no means of guessing whether the student is just being coy in English or whether he doesn't know much about the subject. In short, for general testing purposes, the objective (i.e., structure) section is probably the more useful, but individual differences appear if the composition test is also used.

The following table shows the performance of a fairly large sample of students on both sections of this test.

---

[9] E.g., J. B. Enochs, *Measuring the Ability to Write*. (Unpublished doctor's thesis. University of Chicago) 1946.

WRITING AND STRUCTURE TESTS

	Structure (95 items)	Writing (100 point scale)
Decile	*Score*	*Score*
9th	78	84
8th	67	79
7th	57	73
6th	49	68
5th	44	62
4th	31	58
3rd	22	53
2nd	12	47
1st	4	36
Mean	40.62	61.42

Based on the scores of 197 students in seven centers.

These scores were all made by students nearing the end of their training programs. It is, of course, impossible to judge how many of them will stay on (either by choice or force) for further training in English before taking up other work. The reader will be able to judge for himself what would be desirable by examining the writing test of an average student, which is reproduced below.[10] Fifty per cent of the students tested did worse.

TEXT FOR PART I

DIRECTIONS: Read the following story *carefully*. After five minutes it will be taken away from you, and you will be asked to answer a few questions about it in writing.

---

[10] The writing tests for final recording purposes were all re-read by one member of the staff using the same scale. The differences between his scores and those of other staff members who read the papers were negligible (never exceeding two points) but this technique eliminated even this slight variation. This student's score of 62 is close to the mean of 61 and median of 62.

Two Old Ladies in a Street-Car

It was wintertime and it was very cold. Two old ladies got on a street-car. They sat down near a window. One of them got up from her seat, opened the window, and sat down again. The other one, furious, got up, closed the window, and sat down again.

The first lady: - Conductor! Open the window! I don't want to die from lack of air!

The second lady: - Conductor! Close the window! I don't want to die of pneumonia either!

The poor conductor did not know what to do. He wanted to please both ladies. He wanted to keep them both from dying. But it was impossible; one wanted the window open, the other wanted it closed. How was he to solve such a hard problem?

While the conductor was thinking of a way out, the two ladies started arguing in loud voices. The noise disturbed everybody in the street-car.

"Conductor This business can't go on forever! We must do something!" cried one of the passengers.

"But what can I do, sir?" asked the poor conductor, hoping that somebody would give him advice.

"First open the window, and one of the ladies will die of pneumonia. After that, close the window, and the other one will die from lack of air. Once they are both dead, the street-car can go peacefully on its way."

DIRECTIONS: You are to answer the following questions about the story which you have just read. Write your answers in the blank space below the questions, numbering each answer correspondingly. Write a complete sentence for each answer. You have ten (10) minutes in which to finish Part I.

1. What season of the year was it, and how was the weather?
   It was winter season and the weather was very cold.
2. What did the two ladies do?
   The two ladies went into a street-car.
3. Where did they sit?
   They sit near a window.
4. What did the first lady do?
   She closed the window.
5. And the second?
   The second lady opened the window.

6. Whom did the first lady call, and what did she say?
   She call the conductor in order that he closed the window because she didn't want to die by lack of air.
7. And the second?
   The secon call the conductor too, but she wanted the window close because she was afraid of die by a pneumonia.
8. What did the conductor do?
   The conductor start to thinking of a solution.
9. What did the ladies do while the conductor was thinking?
   The ladies continued arguing in a loud voice.
10. What did a passenger suggest?
    To kill both ladies.

PART II

DIRECTIONS: You are to write five sentences about each of the following themes. Use the blank spaces provided. You will have fifteen (15) minutes to finish Part II.

A. Describe the person and occupation of each member of your family.
   1. My father is a Bank's manager.
   2. My mother takes care of home and works at her free time in my father's business.
   3. I have two brothers and they are studying.
   4. My junger uncle works to a oil company.
   5. My older uncle works to my country's government.

B. Tell in detail what you did last Sunday.
   1. As every Sunday I waked up about eleven thirty.
   2. At noon I went to eat my lunch where I use it.
   3. After I ate come back home and spent about three hours reading a novel.
   4. At four o'clock I took a bath. then I dressed up and went out to the movies.
   5. After show I eat my supper and went home to study until middle night.

To summarize our conclusions from these data and what we have observed, the first is that the standard of English for foreign students should be kept high. In many cases this decision will mean that it will have to be raised. We believe that the evidence here presented, limited in scope though it is, indicates that a higher standard is desirable. For the reasons pointed out at the beginning of this chapter, the establishment of the necessary relation between performance during training and success in later use of the language will always be difficult to secure;[11] but additional data would certainly be desirable.

If the matter still appears debatable, however, it would be better to err on the side of too high a standard rather than too low. Only with such a high standard will we really do justice to foreign students. They are human beings as well as bearers of tuition and grants-in-aid. The students come with a feeling of responsibility. The opportunity for study in America is their "great chance," and they are under considerable pressure to justify the confidence placed in them by their governments, families, or others who have made their coming possible. Coming from different educational systems and having to adjust to a foreign culture, they face enough problems without having too great linguistic difficulty. Of necessity they will labor under linguistic handicap; the best we can do for them is to make certain that it is as slight as possible and to reduce the amount of disappointment and failure by requiring a high level of achievement when they are accepted for study here.

Severe pressures work against such standards. Quick cures for anything are always popular. The student with an opportunity for study in America will be inclined to seize the chance and hope that he can fix up his English in a hurry. Anyone responsible for bringing foreigners to this country will prefer to believe that a couple of month's work here will do the trick and that ample time will then be free for other work.

But any great optimism in this regard seems to us to rest on an undue minimization of the great gap between elementary English and full English. A number of different procedures

[11] Cf. pp. 249ff.

have been developed for teaching English at the elementary level, and often bitter rivalry has sprung up between their advocates. Despite the differences which are usually stressed and defended strenuously, all these procedures follow a generally similar plan.

They consist of the "most useful" vocabulary of 850 to 2,000 words and the "most useful" sentence-patterns and structural features. "Useful" is generally defined as that selection of English vocabulary, structure, or syntax which will enable the student to express the most things — usually, to express everything. In some cases, the passive skills of reading and listening are also considered; in that event frequency of occurrence in English is also taken into account (as well as usefulness or coverage) so that the student will command the words, forms, and constructions which he is most likely to encounter in oral or written form. As Fries and Traver have shown in regard to vocabulary,[12] there is considerable overlapping between the vocabularies of these systems, and their other features show the same similarity.

Since we have mentioned these systems, we may digress for a moment to speak to a question often asked of us in the course of the Investigation. Which of these systems is the best? This is an acute, practical problem for any person charged with elementary English instruction. Shall he adopt this system with its accompanying texts, or that one with its texts? At first glance it might seem that some direct comparison should be possible to determine which of them teaches "elementary English" most effectively.[13] A moment's reflection reveals that such comparative judgments are impossible. Each of these systems defines "elementary English in terms of itself." What it contains is, for it, "elementary English." In spite of the overlapping already mentioned, there are items of vocabulary and everything else which are unique to it. A common measuring device is hardly possible practically. And if one can be made which limits itself to those elements common to all systems, it would be open to objection on theoretical grounds since it would probably be forced to ignore larger portions of many systems than it could include. The systems at this level are incommensurable.[14]

---

[12] C. C. Fries and A. A. Traver, *English Word Lists*. (Washington: American Council on Education, 1940).

[13] For our interests here the question whether some of these systems are intended to serve as an international auxiliary language is irrelevant. We are considering them here only as stepping-stones to full English. Taking up the other question (which lies just outside the scope of this book) merely obfuscates the present question.

For our present purposes, we should not be too interested in this level, for we are considering these systems not as "little languages" but as means to full English. The proof of effectiveness should then be obtained by comparative study at the level of full English, not within the limited elementary system. For this purpose experiment would cover several comparable groups of students. They would begin their study of English by means of one of these elementary systems and move through it to experience with a variety of materials in full English. Then, at the end of some sufficiently long period, the relative achievement of all groups would be measured by a series of common tests, based on full English without regard to the various elementary methods.

Data of this sort have not been accessible to us for a number of reasons. For one, as we have seen in the preceding chapters, the number of language students diminishes in more than geometric proportion to the time of study involved. Our difficulty in securing continuous data of this kind was increased by the war, for many of the programs with the largest number of continuing students in English were situated in foreign countries and were thus made inaccessible or were interrupted by hostilities.

---

[14] The Investigation had access to only one comparative experiment of this kind. The detailed evaluation of it (H. B. Dunkel, "The Basic English Experiment in the Evening Elementary Schools of New York City") can be obtained from the Investigation.

A much more important fact, however, is that there has been relatively little interest in the higher levels of English. While many workers have developed systems of elementary English, few of them have worried about what happens to the student after that. A number of causes probably combine to produce what seems to us this unfortunate result. The decrease in the number of students at the higher levels makes publication of materials difficult. A commercial publisher prefers to aim at the market of millions of beginners rather than the much smaller one of those who stick with the language. Mending broken English and working at second-hand with other people's products is never too interesting professionally. Unless a teacher is in a situation where he can start with a group of beginners and carry them all the way through, he is unlikely to find the work at upper levels very interesting. He will find that the diversity of prior training and the bad habits and wrong ideas (in his opinion, at least) which the students have acquired earlier make any organized planning difficult. It is much easier and more rewarding to start with a clean slate of beginners.

A further cause, which we would particularly emphasize here, is a tendency on the part of the profession to minimize the magnitude of the problem which exists at the upper level. The assertion is frequently made that once the student has laid a solid linguistic foundation, he will have no difficulty in mastering the rest of the language. The proverb says that "well begun is half done," but even it does not claim that the whole task is completed. We have seen no evidence that greater optimism is justified in linguistics. On the contrary, both practical experience and theoretical considerations indicate the immense difficulty of this step from the elementary to the advanced levels.

At the elementary stages, as the similarities of the various systems show, the choice of materials to be covered is fairly easy. As regards vocabulary, for example, a certain limited number of words, roughly two thousand, are going to turn up whatever the student reads; and the application of criteria of selection other than frequency of occurrence also results in clear choices for the vocabulary of elementary work.[15] As was pointed out in Chapter II, this limited "mastery" is worth little in itself. He must add to this knowledge if it is to be much good to him. Let us then examine the problem he faces in making these additions.

To continue speaking purely in terms of vocabulary, there is too much of a tendency to label everything beyond the elementary stages as "technical vocabulary." Presumably the student will pick up the jargon of his field or will already be familiar with it when he comes. Thus the assumption is that, provided he has a basic vocabulary of a few thousand words and a knowledge of the special terms of his field of interest, he will have all that he needs.

The unsoundness of this assumption can easily be demonstrated by a specific example; and since we are all interested in languages, the illustration can be taken from a well-known book on linguistics. A hasty tabulation of page 1 of Bloomfield's *Language* on the basis of the Thorndike-Lorge Word Book[16] gives the following analysis of relative frequencies of the words on that page:

---

[15] Because of greater ease with which vocabulary can be examined and because more work has, consequently, been done with it, we state the issue here in terms of vocabulary. The problems and principles are the same in regard to morphology and syntax; and the discussion should be understood as applying equally to them.

[16] Thorndike, E. L. and Lorge, I., *The Teacher's Word Book of 30,000*

Thorndike-Lorge Ratings	*Frequency by Thousands*	*Occurrences on Page 1 of Bloomfield*
AA	1st thousand	248
A	2nd thousand	27
49-30	3rd thousand	17
29-19	4th thousand	7
18-14	5th thousand	6
13-10	6th thousand	6
9-8	7th thousand	5
7-4	8th-10th thousands	10
3--	Beyond first ten thousand most frequent words	5

Anyone who has ever worked with one of these frequency lists knows how their findings must be interpreted with caution. None the less, they represent some of the most exact data we have; and the point at issue is still clear even if we interpret the table most cautiously.

The average foreign student coming to this country apparently does not have a vocabulary of much over 2,000 words. But to allow for sampling-errors in the list, words known through cognates, and those derivable from better known members of the same word-families, we can take as a critical point in the preceding table the 5,000 most frequent words, and we will assume that the student knows all these. But the table demonstrates that even a student with this large a vocabulary will find that many (26) words on the first page of Bloomfield are unknown to him.

Are all those "technical" words which the student of linguistics will pick up in his reading, or are they words which he will encounter only in linguistic materials and which he cannot be expected to know unless he is a specialist

---

Words. (New York: Columbia University Press, 1946). The conversion of the ratings into frequency-levels by thousands is that suggested in the *Word Book*.

in linguistics? Some of them can be so classed: e.g., *linguistic* itself (Thorndike-Lorge rating = 1), *predicate* (1), *dictionary* (9). But what shall we say of *rigid* (11), *conventional(ly)* (9), *reasoning* (13), *masquerades* (5), *sophisticated* (4), *medieval* (9), and *comprehensive* (7)? Certainly they are not technical words of linguistics or anything else. They are part of the common stock of any educated English writer.

Since it may be argued that the first page of any book is not typical, a second random choice was made, page 101. This falls in the midst of the discussion of the phoneme, and scarcely any other topic is more typical of the school of linguistics identified with Bloomfield. The analysis of page 101 must be on a slightly different basis because of the technical material:

Thorndike-Lorge Ratings	*Frequency by Thousands*	*Occurrences on Page 101 of Bloomfield*
AA	1st thousand	232
A	2nd thousand	33
49-30	3rd thousand	7
29-19	4th thousand	0
18-14	5th thousand	4
13-10	6th thousand	4
9-8	7th thousand	4
7-4	8th-10th thousands	4
3--	Beyond 10,000 most frequent	7
Proper names		22
Technical		45

In the preceding table, no attempt was made to classify in terms of frequency either the forty-five obviously technical terms like *spirant, unvoiced, lateral, stop,* and *nasal.* The proper nouns (names of languages, language groups, and countries) were also omitted. Both these classes of words are those which the foreign student of linguistics is likely to know or to add quickly to his vocabulary. Such is not the case with the other words at the end of the list (again with the Thorndike-Lorge ratings in parentheses following them) such as: *shin* (5), *protrude* (4), *eddying* (8), *variant* (1), *customary* (12), *preceding* (9), *trill* (3), *friction* (9), and a number of others. This selection illustrates the point as well as the other. In addition to his vocabulary of basic words and technical terms, the student will need a sizable vocabulary of words which are neither "elementary" nor yet "technical."

We are, of course, merely using Bloomfield's own words as an illustration of the point he himself explicitly makes:

> The myths about peasants, workingmen, or savages who use only a few thousand words have no foundation in fact; in so far as one can count words (ignoring, for instance, the inflected forms of a language like ours), every adult speaker uses at least somewhere around 20,000 to 30,000 words; *if he is educated-- that is, if he knows techincal and learned words-- he uses many more.*[17]

Yet it is with educated speakers that the foreign student will be in contact. If he has "mastered" English only at the elementary level — i.e., if he has a vocabulary of only 5,000 words or fewer and has a similarly restricted knowledge of morphology, syntax, and idiom -- he will find himself, in terms of vocabulary, for example, at least 15,000 words short! And as may be seen from Bloomfield's statement, this estimate probably errs on the side of modesty. The student (and those who claim to help him) must, therefore, find some way of getting into his vocabulary these additional words which the educated native speaker and writer will use. Dictionaries will be of limited help in reading, but even the most prolonged practice in thumbing them will not enable him to keep up with the speech of an American friend or lecturer.

---

[17] *Language*, p. 277 (the italics are ours.) In view of the generally accepted belief that in writing one uses a larger and more varied vocabulary than in speaking, it is worth noting that Bloomfield refers to the "speaker."

For an account of past studies and further evidence, see R. H. Seashore and L. D. Eckerson, "The Measurement of Individual Differences in General English Vocabularies," *Journal of Educational Psychology* XXXI (1940) 14-38. See also G. W. Hartmann, "Further Evidence on the Unexpected Large Size of Recognition Vocabularies Among College Students," *ibid.*, XXXVII (1946) 436-9.

Yet these words are exactly the ones which the student has the most difficulty in learning, simply because they appear with such relative infrequency. Of course one may hope that he will pick them up eventually. This is true of all his English; he could have, from the outset, stumbled around for himself. The educational question is whether there is some way which will help him do it more efficiently. And work at this level is much harder for both teacher and student.

Sometimes classified word lists have been suggested to overcome this difficulty; i.e., a lesson on "Parts of the Body" gives *ankle, shin, collarbone,* and the rest. If these masses of words are sufficiently drilled at the time and adequately reviewed later (conditions usually not met), this method works.[18] Another way of providing original exposure to the words or giving the needed review (so that the masses are not merely learned once and then forgotten) would be readings designed to give abnormally high concentrations of these rarer words. Instead of "grading materials down" to the vocabulary of the 2,000 most frequent words, the aim would be to "write them up" to the level of the 5,000 - 10,000 range as far as possible—though a heavy loading of the common structural words would naturally be inevitable. Only by such artificial concentration can the student be brought into contact with these less common words often enough to retain them. The result might sound like a bad parody of the high-flown style; but, like the passages "written down," the purpose of these readings would be pedagogical and not literary. In elementary materials, repetition has been constantly

[18] A good selection of the type of word involved in this advanced vocabulary can be found in Pankey and Sachs, *Five Thousand Useful Words*, (Nashville: Southwestern Co., 1936).

emphasized; yet the basic words are those which, by the nature of language, are most likely to be repeated. We clearly need even greater effort devoted to the less common and less fundamental words which require it still more.

Before leaving this point, we should repeat the warning given earlier. This discussion has been in terms of vocabulary — and vocabulary for reading. This is not the whole of language. We have used it for the example only because the basic data (the frequency counts of written materials) were available and easily usable. But an analysis of vocabulary for aural comprehension or oral command would necessarily show the same situation, and syntax and idiom are equally important. There has been a striking neglect of this higher level of English.[19]

In sum, the two urgent needs of instruction in English as a foreign language for foreign students studying in the country are: *(a)* a raising of the general standard required and *(b)* the preparation of materials which will enable students to cover this higher level of language more adequately and efficiently and thus to meet this higher standard. We believe that the English examinations prepared by the College Entrance Examination Board and to be administered *before* the student comes to this country are definitely a step in the right direction. If American institutions demand a sufficiently high level of performance on these examinations, further training in English here can then deal with the advanced level where much help is needed.

---

[19] We Know of only one book (Kelly, *An Advanced English Course for Foreign Students*) which attempts to grapple with these problems; but it is not adapted to American use and has certain grave defects.

UNIVERSITY OF CHICAGO LANGUAGE INVESTIGATION
TESTS FOR ENGLISH AS A FOREIGN LANGUAGE

*Reliabilities*① *and Intercorrelations*

	Reading	Aural	Structure	Writing	Elementary Structure⑤
Reading	$\boxed{.90}$	N=102	N=156	N=155	N=100
Aural	.75	$\boxed{.81}$	N=107	N=107	N= 28
Structure	.70	.66	$\boxed{.96}$	N=190	N= 56
Writing	.58	.57	.70③	$\boxed{X}$	N= 49
Elementary Structure	.62②	.39	.74	.77④	$\boxed{.97}$

①By split-half corrected by Spearman-Brown. The reliabilities were based on the following *N*'s: Reading 292, Aural 155, Structure 339, Elementary Structure 161.

②The correlation ratios are .62 and .77, the latter of which (reading on structure) is significantly greater than *r* at the 1% level.

③The correlation ratios are .73 and .86, the latter of which (structure on writing) is significantly greater than *r* at the 1% level.

④The corrplation ratios both round to .87 and are significantly greater than *r* at the 1% level.

⑤Because few students who took Elementary Structure also took any of the advanced tests, all correlations with this test are based on small *N*'s.

Chapter VIII

A Statement of Conclusions and Further Needs

The preceding chapters have given detailed descriptions of our procedures and of the various programs with which we worked. The varied aims, methodologies, situations, and results which characterize these individual language courses – particularly the experimental undertakings described in Chapter VI – make generalization hazardous. Yet the single trees, considered collectively, do form a forest; and language teachers are probably more interested in a general picture than they are in the nature and success of any specific experiment.

The controversies about the newer programs[1] produced heated debate on such topics as the actual achievement of the experimental courses, the superiority (if any) of their results over those obtained by more conventional approaches, the efficacy of certain features of them (intensive drill, initial aural-oral emphasis), and the like. Whatever may be the shortcomings of our evidence (and we have gone to considerable pains in the preceding pages to indicate these weaknesses as clearly as possible), our data constitute at present the most extensive and objective answers to some of these questions.

---

[1] The statements pro and con about these and other issues (which date from the period of the ASTP) are fully recorded in P. F. Angiolillo, *Armed Forces' Foreign Language Teaching* (New York: F. S. Vanni, 1947).

In addition to judgments about the present state of affairs, however, we feel some responsibility for stating what appear to us to be the most pressing needs of language study in the near future. Obviously our hypotheses about these matters are not proved by the data but are only suggested by them and by our experience in observing and testing a fairly wide variety of language courses.

Even the conclusions must be more tentative than would be desirable. In Chapter I we pointed out that an investigation should develop certain hypotheses which can then be tested in controlled situations until they merit the name of demonstrated conclusions. As we also pointed out in Chapter I, existing circumstances did not permit this procedure. None the less, we have used all possible caution and believe that the following summaries and conclusions constitute the soundest judgments which can be made at the present time.

DESCRIPTION OF THE EXPERIMENTAL COURSES

Objectives

All the programs had essentially the same set of objectives: ability to speak, to understand, and to read the foreign language. The initial goal was the development of oral and aural skills, and the ultimate goal was the attainment of reading proficiency. Functional grammar, intended to contribute to both oral-aural and reading competence, was emphasized in all phases of many courses, but in few if any programs was the active manipulation of structural patterns an end in itself. Translation from English into the foreign language according to grammatical rule was generally not practiced. There was virtually no insistence on ability to write the foreign language.

Rationale

The basic assumption on which the programs rested was that a second language, like a first (i.e., native) language, is most naturally acquired in its spoken form and that a solid and lasting ability to read a language best results from previous acquisition of that speech which its written form symbolizes. Thus the ear and tongue are to be trained first, and the eye only later. This became the justification of teaching command of the spoken language even where, within the liberal-arts tradition, the cultural reading aim could and must not be forsaken. This distinguished the program, in the last analysis, from a "Berlitz-type" course.

A complementary and also partly justificatory assumption was that in the world of today, struggling as it is toward a single peaceful society, a nation must avoid linguistic isolation by training more of its citizens to use the languages of other nations, and so equipping them to take an intelligent part in international affairs.

A third assumption was that most American students are more highly motivated to elect foreign-language study and to persevere in it, if the reward of their choice and their effort is to be proficiency in using the language in its spoken form, than if their only accomplishment is to be reading ability.

Given these three assumptions on which recourse to an initial oral-aural approach was based, we may proceed to examine the principles of oral-aural instruction as followed in the experimental programs.

The cardinal principle was that the acquisition of oral and aural ability, in a second as in the first language, is primarily a mechanical rather than an intellectual process, comparable to a skill in art or craft and not to a mental discipline. Its first requisite is *practice* – practice in hearing and imitating speech utterances, and quickly associating them with their meaning while doing so. This practice consists of simple mimicry-memorization at first, during the formative period of the sound-producing habits; later it consists in increasingly varying the basic patterns by substituting different forms or different words for one or more of those contained in the original utterance, and by answering questions; eventually the already acquired patterns are used and varied in real or simulated conversation. The more hearing, imitating, repeating, memorizing and eventual varying of the utterances the better, until the point is reached where they are *over-learned*. Ideally such practice should go on continually; and in conformity with this principle, experimental oral-aural courses required the maximum possible amount of the student's time in which the practice could be directed and supervised by native speakers of the language. Furthermore, each student must be provided with opportunity for *individual* practice. Choral practice is valuable up to a point, for straight listening and mimicking, but pattern-variations and actual conversational interchange must be on an individual basis; consequently in oral-aural experiments the number of students participating in classroom practice was held to the absolute minimum possible.

While requirement number one was practice, which is mechanical, a second essential of the oral-aural approach was analysis, which is intellectual. The adult learning to use a second language differs from a child learning his native tongue in that he can, theoretically at least, bring his intellect to bear on his problem and can speed up his learning process immeasurably through generalizations, shortcuts and insights into the way the language operates, if and when he understands its structure analytically. In accordance with this principle experimental courses provided, along with practice, instruction in the phonology, morphology, and syntax of the language. In this connection, a second principle applied in some programs was that, in order to be meaningful, grammatical interpretation and analysis must be made strictly in terms of patterns already familiar to the learners through the mimicry-memorization process.

Organization

In conformity with these complementary principles of mechanical and analytical learning, two kinds of instructional situation were provided: (1) "lectures" or "demonstrations" by the linguist in charge, before the entire student group (or before sections of fifty or sixty students if the enrollment was large), and (2) "drill" or "practice" sessions, limited to ten or twelve students each and conducted by native speakers of the language.

The number of contact hours per week, embracing both lectures and drill sessions, varied between six and fifteen, the commonest being eight. Generally, two different interpretations prevailed with respect to the equating of these contact hours with semester hours or credits. Some programs were designed to bring the students to two-year levels of proficiency in one year; in these the amount of credit traditionally given for regular language courses was doubled, and the students were expected to devote additional time to home study of the language over and above their weekly class hours. Other programs, while roughly doubling the weekly contact time of conventional courses, still aimed to cover but a single year's work in a year and regarded the increased classroom practice as a partial or total substitute for outside study; in such courses credit hours remained the same as for conventional courses.

Staff

In general, the teaching staff of an experimental program consisted of a course director, who was a regular member of the foreign-language department of the institution, and of instructors variously designated as "drill masters," "guides," "informants," or the like. The course director gave the lectures, and planned and supervised the work of the instructors; the latter conducted the drill sessions.

The instructors were for the most part native speakers of the language in question, but their background and experience as foreign-language pedagogues varied considerably, in some cases even within a single program. In some programs it was held that native speech was the only requisite of a drill instructor, and that indeed previous experience in teaching a language might lead him to spend practice time on grammatical or other professorial digressions when he should be serving as a model or making the students use the language during every moment of the precious contact time. Where this attitude prevailed, drill instructors were recruited from any available source – from foreign students attending the institution to purely chance members of the community. They were expected to serve short terms of one or two years only. In other programs it was felt that the native-speaking instructors should also be experienced teachers who as persons of education and culture could impart to the students a first-hand acquaintance with the people and institutions of their native countries. In small-scale operations, native instructors of the latter type were generally available within the regular teaching personnel of the language department; in the larger programs both types of native speaker – the untrained recruit and the experienced pedagogue who was a regular member of the staff – were likely to be serving side by side.

Materials

The most widely adopted series of textbooks specifically designed for use in oral-aural courses were the manuals of *Spoken French, Spoken German*, etc., first developed for use in Armed Service programs and later made available for civilian use.[2] With the material organized according to an identical plan for all languages, these manuals present thirty chapters or "units" consisting of "basic sentences" to be mimicked and memorized; "hints on pronunciation"; a section of "word study" designed for structural analysis of the basic material and organized by grammatical categories; a "listening in" section consisting of variations of the basic sentences, designed primarily for aural practice; a set of exercises calling for completions, substitutions, choices and the like entitled "What would you say?"; and finally suggestions for carrying on "conversation." The units are topical, and carry titles such as "What's Your Trade", "Laying in Supplies", "Fill 'er Up", etc. Each of these topical units contains material for six to eight hours of class work, and directions are provided for individual home study and review. In every text the foreign-language material does not appear in normal orthography during the first six units, being reproduced instead in phonetic transcriptions designated as "aids to listening."

---

[2] French, Italian, Portugese, and Spanish are published by Heath, and the other languages of the series by Holt.

Other texts extensively used were Harris and Lévêque, *Conversational French* (Holt); Rehder and Twaddell, *Conversational German* (Holt); La Grone, *Conversational Spanish for Beginners* (Holt); and the Decca Language Series entitled *Les Aventures de Robert Martin, Las Aventuras de Roberto Martín,* etc. All these texts are organized primarily by topics of everyday experience, and provide basic material for mimicry-memorization together with grammatical analysis and practice exercises. All of them introduce conventional orthography from the beginning; and do not make use of phonetic transcriptions.

No short list of reading texts would be representative of those selected for the experimental programs. Adoptions varied as widely as in conventional courses.

Audio-Visual Aids

Though there has been much talk about the value of various electronic sound-machines in foreign-language instruction, little if any use of such devices was observed even in the sporadic cases where a language department owned one. Course directors did not seem yet to have worked out practical techniques for using wire or tape recorders, which require deft operation and carefully planned integration with course procedure if they are to prove efficient, time-saving substitutes for, or supplements to, living models and not merely time-wasting curiosities.

Here and there textual materials and supplementary exercises were locally recorded on phonograph disks and utilized for aural practice, and occasional recordings were made of students' speech. In general, however, the only audial aid extensively used was the ordinary phonograph. Records are available, for example, reproducing the basic sentences and aural exercises of the Heath and Holt spoken language texts, and in several programs these records were used by the linguists in charge for demonstrating the sounds of the language, or by the students themselves for extra practice in hearing and repeating the lesson material.

ACHIEVEMENT IN THE EXPERIMENTAL PROGRAM

The results achieved in different programs were far from uniform. The present over-all appraisal is a generalization based on the average level of achievement, and only occasionally calls attention to important and clearly explicable deviations from that mean. It will further be understood that we are talking in terms of the average student within a given program. On our measures the deviations from the mean performance are usually not great in relation to the full range of possible measurement even in the cases of the "best" and the "poorest" students. In other words, even exceptional individuals need not be thought of as achieving perfection on the one hand, or zero on the other.

Oral Production

For the reasons given in Chapter IV (pp. 97-98), we failed to secure extensive objective evidence on students' oral competence, and the following appraisals are based in large part on our personal judgments.

As regards pronunciation, short utterances (ten to twelve syllables) could be immediately imitated by most students almost perfectly. Enough time was never available, however, for sufficient drill in this direct imitation to produce fundamental changes in the students' speech habits. We believe that fifty to a hundred times as many repetitions as were actually feasible would have been required to inculcate new habits of articulation and intonation. As it was, students tended to relapse rather quickly into the phonemic and intonational patterns of their native tongue. An utterance which had originally been imitated very accurately often reappeared later with a partial or complete substitution of English phonemes. This tendency to revert to native sounds increased sharply when the utterance was long or otherwise difficult, and the sight of writing (even a phonetic transcription) had the same effect. Another factor which apparently retarded or even halted some students' development in pronunciation was the urge to avoid conspicuous excellence.

In many programs, lectures given early in the course dealt with phonetic problems and described the phonology of the foreign language, on the theory that students can improve their speech through conscious understanding of the physiology of sound production as well as through direct imitation. Clearly the actual contribution of phonetic analysis to speech improvement can be measured only under more controlled conditions than we encountered in any program. One need only observe an average group studying descriptive phonetics to conclude that, while some students do seem to benefit from it, to many students it means less than nothing.

Nevertheless, despite the failure of most students to retain the excellent pronunciation observable in immediate imitation, the majority were observed to have a continuing production superior to that of conventional students who were taught to speak by rules of "how the letters are pronounced," and who have never heard a native speaker. Also, students entering an oral-aural course after previous study of the language, in high schools where English habits were tolerated, generally never succeeded in matching the pronunciation of those who were taught from the beginning to imitate native models.

We saw few exhibitions of what could honestly be called spontaneously fluent speech. While many students could participate in memorized conversations speedily and effortlessly, hardly any could produce at length fluent variations from the basic material, and none could talk on unrehearsed topics without constant and painful hesitation. The main contributing factors seem to have been slow vocabulary recall and difficulty in manipulating grammatical forms.

The range of vocabulary and idiom was strictly delimited by the topics and situations treated in the conversational textbooks studied. As noted above (see p. 33), most such textbooks are heavily loaded with specific environmental terms, and in compensation omit many common words necessary to a minimum adequate vocabulary. Even within the range set by the textbook, automatic recall of any but the most frequent words was rare. After reading had been introduced into a course, new words and expressions met with in this connection were generally slow to find their way into the students' active vocabulary.

Regardless of the students' understanding of the structure of the basic material which they mimicked and memorized (i.e,, their passive grammatical knowledge), they had little ability to manipulate forms and constructions actively in their attempts at variations or at free speech. Knowledge of linguistic structure enables one to "work out" a desired pattern if given enough leisure, but the structure itself is so complex that it cannot be manipulated consciously in speech production without endless practice in minimal variations of form and construct – and in no program was time available for quantities of such drill.

Aural Comprehension

Generally speaking, students' aural comprehension as observed in class work was better than their oral production. Most of them learned adequately to comprehend the basic material of their conversational textbooks as spoken normally by their instructors, as well as variations of that material designed for aural practice. They were not, however, quick to understand casual or unexpected utterances for which they were not prepared. Furthermore, their comprehension was of course subject to the limitations of the relatively narrow and specific vocabulary of their lesson materials.

The objective data on aural comprehension yielded by our testing program have been extensively analyzed and interpreted in foregoing chapters, so that the briefest statement of findings will suffice at this point. By and large, experimental students failed to understand the phonographically recorded utterances of an unfamiliar native speaker, delivering unfamiliar though easy material, significantly better than did conventional students. It is true that the causes of this similarity of performance were seemingly different for the two types of course. While conventional students lacked aural ability because they had not been taught it, experimental students failed of higher attainment in part because the limits of their comprehension experience (particularly in terms of known vocabulary) were those of their textual materials and were much narrower than those within which the tests were designed to measure aural competence. With due allowances made also for the shortcomings of the tests, the fact remains that most experimental students did not give evidence of approaching native aural ability.

Reading and Vocabulary

We have no objective data on the levels of reading proficiency eventually reached by students who progressed from experimental courses into advanced reading or literature courses. Within the short period of our project, too few students continued into advanced work at the same school. Our testing program covered experimental groups at the end of one and two (rarely three) years, and provided comparisons of reading achievement with conventional groups at those same stages. In general, as shown in the analysis of test data in Chapter VI, the experimental groups did not demonstrate so high a level of reading proficiency at the end of their first or second year as did conventional students. In a number of cases the results were significantly poorer. There appear to have been two chief reasons for this outcome: (1) in many programs so much time was devoted to the initial oral-aural objective that much less reading was done than in conventional courses; (2) as some programs were organized, the responsibility for teaching reading devolved upon drill instructors who were unequipped to deal with the very different problems of developing reading ability. These instructors had little knowledge of what particular constructions constitute hazards for the English-speaking student, or what vocabulary items should be emphasized as worthy of inclusion within his active or passive range of control.

In the vocabulary parts of the reading tests, the experimental students knew fewer words than conventional groups. The reasons were presumably the same as in the case of reading, and the specific vocabulary items learned from the spoken-language textbooks did not coincide to an appreciable extent with high-frequency items used in the reading tests.

On the grammar sections of the Co-operative tests, experimental students generally tended to be as successful as conventional groups. As contrasted with their comparative failure in reading and vocabulary, their grammatical proficiency seems attributable to the early and continuous study of structure in connection with their oral-aural materials, under course directors who were in substantially all cases experienced teachers of linguistic analysis.

VINDICATION OF ASSUMPTIONS

The following two assumptions underlay the experimental courses: (1) that oral-aural proficiency is the best gateway to reading ability; (2) that students are highly motivated to study a language via the oral-aural approach.

1. As far as we were able to follow the experimental groups, i.e., mainly through the first and second years of study, the evidence from the tests indicates that superior reading skills were developed in those programs where reading received the greatest time and emphasis. The highest levels were reached in reading-method courses which featured moderate amounts of classroom oral-aural practice directly related to the material read. On the other hand, experimental students whose oral-aural command was limited to the scope of their spoken-language textbooks were observed to be unable either to understand reading material presented to them *viva voce*, or to discuss in the language the content of their readings.

As to the assumption made in some quarters that oral-aural competence *automatically* creates reading ability and that consequently the latter need not be specifically taught, there is evidence *per contra* so far as the experiments observed are concerned. It may well be that a hypothetical student who has acquired near-native oral-aural competence (including the possession of, say, a twenty-thousand-word active vocabulary) would find himself able to read the language excellently without specific training in that direction. However, this hypothesis is hardly applicable to the students observed in experimental courses. There, only those who were carefully and expertly taught reading skill as an attainment distinct from that kind of oral-aural competence which proved to be teachable within the scope of the program, were able to demonstrate notable success in reading. Finally, low correlations of reading and aural test results from both experimental and conventional programs furnish conclusive evidence that (at least as sub-native levels of competence) oral-aural and reading proficiency constitute separate, independent skills which do not develop one from the other but rather only from direct training in each separately.

2. With few exceptions language students claim to be more highly motivated by oral-aural than by grammar and reading goals. Many elect the experimental courses eagerly with the purpose of getting a speaking knowledge. As the course begins they are pleased to find themselves able to say and understand the simple phrases of Lesson One. By Lesson Five they are convinced that if they went to the foreign country they could "get by." But as drill-sessions succeed each other with relentless regularity, and as the material becomes more and more difficult to assimilate, many begin to tire or to lose interest and fail to apply themselves to the extent the method demands in order to assure continuing success. Their home study, in which the necessary over-learning could be achieved, is neglected, memorization does not keep pace with mimicry, the tempo of drill sessions is slowed, and boredom sets in. In short, while the average student professes continuing enthusiasm for the goals and the method, in reality his dwindling application is a factor which operates in preventing average outcomes from reaching higher levels.

ADHERENCE TO PRINCIPLES

To what extent, in actual practice, did the experimental programs adhere to the principles on which they were based?

The prevailing number of contact hours per week, although it was invariably the maximum possible within the total course of study at a given institution, did not provide enough time for that repeated drill with the same material which alone results in the necessary over-learning, or for the protracted practice with pattern variations which alone leads to flexibility and ultimately to fluency. Furthermore it was not possible to require sufficient amounts of laboratory or home study, involving unsupervised drill with records or the like, to compensate for the shortage of classroom practice.

The number of students participating in drill sessions was often not rigidly controlled, especially in large programs where even with numerous drill-sessions scheduled at different hours of the day, the inevitable conflicts in students' individual course schedules frequently necessitated the unbalancing of drill-sections; a consequence, of course, was that some students received less individual attention and practice than others.

In the grammatical analysis phase, the principle of keying the lectures to the drill was frequently interfered with by schedules which had students who attended the same lectures drilling a given set of material at different times, some before the lecture and others after it.

There were very few exceptions to the practice of employing native speakers for drill instructors, and those who did not so qualify were nearly bilingual or had entirely authentic pronunciation.

FURTHER NEEDS

The experimental courses evaluated by the Investigation generally failed to produce near-native oral-aural or reading proficiency in the American student of a second language in one or two years. In other words, within the total instructional time available for these experiments, the newer procedures and techniques have not proved themselves impressively in training students of average aptitude and motivation. Perhaps the goal was unrealistically high, but it was that sought by many experimenters; and only a near-native ability will justify some of the advantages claimed for the skills. In any event, it is quite improbable that within the normal high-school or college curriculum still more time can be obtained than has been available to current experiments. Therefore, if the newer programs are ultimately to vindicate the assumptions which underlie them, improvements in procedure and technique must be brought about within the organizational framework (including the time allowance) already acquired. It seems to us that such improvements can and will be effected through careful attention to a number of needs which suggest themselves rather clearly.

The Need For More Precise Definitions

Despite eager discussion and experimentation in the fields of aural comprehension and oral production, language teachers are still working their way toward generally accepted definitions of these abilities. Since the level of native competence will obviously not be reached soon, some explicit standard short of that must be set up, at least temporarily. Thus far there has developed little general agreement about such obvious matters as the kind of topic which the student is expected to be able to handle orally and aurally, the total number and range of these topics, the nature and extent of the vocabulary needed, and similar points. The textbooks constitute one approach to evolving this canon for the aural and oral skills; but, as we have seen, they still reflect primarily the personal and somewhat random choice of their respective authors and not a clear body of theory. Standardized tests represent another approach to the problem. In any event, whatever materials help to clarify the issues, the profession needs to develop a more explicit and justifiable consensus on all these matters. Only then can the verbal symbols like "oral command" be meaningful in our discussions, and only then can tests be constructed which will accurately measure the skills precisely as they have been defined and thus demonstrate the students' achievements unequivocally.

The Need to Re-examine the Language Requirement

It is not within our province here to urge the extension, retention, or abolition of the language requirement. The present evidence does lead us to suggest, however, that language teachers might well re-examine the desirability of the requirement in view of the high level of linguistic proficiency they hope to impart to students. The language skills apart from their concomitant values are rewarding only after they have been attained to a considerable degree. Certainly the uses and values commonly claimed for second-language study presuppose a fairly high degree of mastery. The question is, then, whether language teachers can hope to give this amount of skill to the masses of students (many of them possessing slight ability and still less motivation) brought into elementary and intermediate classes by the language requirement. Their presence acts as a brake on the progress of the more promising students. Special sectioning, while it has proved impossible in most small experimental programs, may ultimately avoid this difficulty. But there still remains the problem of bringing this less likely raw material up to the required standard. Though we have spoken specifically of the achievement of only the upper half and upper fifth of the conventional groups we studied, the reader can find in the complete norm tables the present status of the lower half or similar groups. These students languish in the limbo of those who cannot read nor speak nor comprehend. For the profession to contract to deliver linguistic skill to these students is a responsibility not to be lightly undertaken.

The Need for Greater Modesty of Claims

Teachers of all second languages will ultimately benefit from greater restraint in claiming what they can accomplish and from a candid acknowledgement of the limited skill which the average student will acquire within the time of the average course. Several champions of the cause of language teaching have already warned that, whatever improvements the new procedures may eventually make possible, they do not work miracles; but the natural flush of enthusiasm attendant upon new undertakings has raised some overly sanguine hopes. Proponents of the newer methods have made much of the disillusionment experienced by earlier generations of students who, after several years of foreign language work, still felt like deaf mutes. Yet the indignation of students who have been led to believe that they can achieve aural-oral "mastery," but who later find that they lack complete competence, will be even more righteous and will recoil even more sharply upon the profession. Language learning remains a long, hard road, and the average student, despite possible improvements in teaching procedures, will probably not get very near the end of it. A clearer understanding of what is expected of him and of what, in turn, he can expect will produce a better situation in the long run.

The Need for Better Articulation of the Skills

For some time many language teachers have doubted whether initial reading ability could later be easily expanded to include aural and oral commands. On the other hand, some proponents of the oral-aural introduction had assumed that these abilities would lead, almost automatically, to satisfactory reading ability. Many parts of our data cast grave doubt on this assumption. At least under present circumstances (though these are not necessarily immutable, as we shall note in a moment), the various skills must receive considerable individual attention. Course directors who have recognized this fact, however, have confronted a whole series of problems. In shifting from oral-aural work to reading, they have searched, almost in vain, for materials which would effect the transfer with the minimum break in continuity. Another problem has been that of deciding when the transition should be made. Some courses using the aural-oral approach have introduced reading materials after the first week or two; others have contained no written materials other than phonemic transcriptions until well into the third academic quarter. Most course directors and many outside critics have questioned the soundness of whatever particular procedures were used; and many observers would explain some experimental programs' lack of success chiefly on the grounds that the change in emphasis was not made in the proper way, at the proper time, or with the proper preparation. Both these opinions and the evidence suggest that the satisfactory articulation of work with the different skills remains a major problem yet to be solved conclusively by the newer courses.

One aspect of this same problem even involves schedules. Most oral-aural courses have sought as intensive a program as was possible in ordinary academic schedules. Increased class time met an important requisite for the acquisition of the oral-aural skills, the opportunity for extended practice. An analogous need exists in regard to reading; but for it, instead of additional class hours, the student needs more outside reading time. Most programs, however, continued their larger amounts of class contact and thus reduced the time available for reading practice even after that skill had been taken up. This difficulty admirably illustrates the queer guises under which the problems of articulation appear.

Thus far we have discussed the problem of articulation as it now exists. The greatest need, however, is to reduce its present size. As we have seen, existing materials, vocabularies, and nearly everything else connected with speaking and listening differ drastically from those used in connection with reading; and it is this difference, of course, which produces the problem. In our opinion, the gap could be considerably narrowed by attempting to bridge it from both sides. New reading texts embodying fictional and (especially) nonfictional material could utilize subjects, styles, and vocabularies much closer to those usable in oral-aural work than they now do. Materials giving cultural information about a linguistic area, for example, might well be prepared in this fashion. On the other hand oral-aural materials now represent a rather arbitrary selection of vocabulary and content. In their effort to give the specific vocabulary for certain selected situations, they often make some rather startling omissions of words which are basic to the language, whether written or spoken. Less emphasis on the tourist and a greater realization of the wider extent of possible oral-aural situations can bring about greater *rapprochement* from this side too.

The Need for Better Tests

As we have frequently pointed out, markedly better tests for the aural and oral skills will be possible only after these abilities have been defined more exactly. When this clarification has been achieved, then tests for general use can be constructed which will measure validly the elements of the skills as thus defined. Meantime, better and greater use can be made of existing instruments. Without some sort of objective standard, wishful thinking and other influences will continue to affect estimates of achievement.

The Need for Qualified Drill Instructors

If teaching staffs are to include instructors whose primary qualification is that they are native speakers and whose role does not extend beyond supervision of drill practice, such native instructors need thorough orientation as well as continuous training by their directors. They need to have a precise conception of both the potentialities and the limitations of their function. On the one hand, they must be adept at handling groups of American students sympathetically and tactfully while at the same time stimulating them to put forth their best efforts in drill and conversation. On the other hand, they must regard themselves as models of their native spoken language and not as philologists or grammarians; they must remember to teach the language, not to teach about the language. And since all native speakers are not by definition persons of sufficient intelligence and adaptability to learn the role of drill instructor, methods of procurement must be such as to assure a ready supply of those who are potentially qualified.

If, instead of (or in addition to) the type of drill instructor we have just mentioned, trained teachers regularly on the staff are to conduct drill practice in addition to their other assignments, they too need training in the special technique of *drilling* as distinguished from the regular classroom *teaching* to which they are accustomed.

The Need for Improved Use of Audio-Visual Aids and Extra-Curricular Activities

These seemingly dissimilar topics are united by common implications for the improvement of aural-oral programs. Only a full utilization of all possible resources seems likely to produce a satisfactory level of student achievement within the time available. While both audio-visual aids and extra-curricular activities now exist on most campuses, neither of them as yet appears to be making its maximum contribution.

The great difficulty with audio-visual aids is that they have usually not been adequately integrated with the work of the course. Audial aids in the classroom can relieve the drill instructors of the more mechanical aspects of their work and thus free them for more personalized attention to individual students. Outside the classroom, these aids will have even greater value, provided they can be constantly and easily made accessible to students for additional practice. Improved use of audio-visual aids will probably depend on a number of factors: improved equipment, greater variety and better quality of materials, and greater experience on the part of teachers in using these devices. Even under present conditions, however, they are a potential ally whose services are largely neglected.

Language clubs, language houses, and language tables are treated by many staffs in a rather perfunctory manner or regarded as an unpleasant duty. They offer, however, a further opportunity for that practice which is so important in the acquisition of the aural-oral skills and especially in understanding and using variations on those basic patterns which students have mastered in class. Enterprising students can be counted on to organize these extra-curricular activities if they are given the support and encouragement of the language staffs. When these undertakings are well organized and well co-ordinated with regular classroom work, they can serve as effective supplements to it.

The Need for More Precisely Planned Experiments

As we have pointed out, current experimental programs have left unsolved a number of basic problems of language learning. Since language is an intricate mechanism which operates in nearly all human thinking and doing, the principles of language teaching and learning become highly complicated. The second volume of this report, in examining what we now know, will reveal some major areas of ignorance. There is still a serious need for experiments designed to test one by one in carefully controlled situations these fundamental problems which confront the language teacher in his search for perfection. Because of the wide range of factors which must be taken into account, language experimentation is extremely difficult and demands the most careful planning and execution.

Unfortunately, in practice most experiments are hastily conceived, short-lived, and executed under appalling difficulties. Certainly the profession cannot effect major improvements in many of these respects overnight. None the less, greater effort must be made in that direction if results even remotely approaching conclusiveness are to be obtained. The data and suggestions offered in the second volume are intended to help along these lines.

The millennium in language teaching has not yet arrived; but in this field, as everywhere else, men of good will can only continue and redouble their efforts to bring it about.

Appendix A

The Norms for the Aural Tests and for the Russian R-V-G Test*

---

*The basis on which the norm-groups were classified is explained in the text, pp. 64 and 74.

FRENCH AURAL COMPREHENSION NORMS
LOWER LEVEL FORM A

Group OA
(28 students in 2 schools)

DECILES	PT. I (25)	PT. II (30)	PT. III (41)	TOTAL (96)
9	9	13	16	35
8	8	13	13	27
7	4	10	11	23
6	3	9	9	19
5	2	7	5	16
4	1	6	4	14
3	0	5	3	11
2	0	2	2	10
1	0	0	1	5
Means	3.43	7.43	7.22	17.94

Group OB
(305 students in 8 schools)

DECILES	PT. I (25)	PT. II (30)	PT. III (41)	TOTAL (96)
9	7	18	14	33
8	6	14	11	27
7	5	11	9	23
6	4	10	7	20
5	4	9	6	17
4	2	6	5	15
3	1	6	3	13
2	0	4	2	10
1	0	2	1	7
Means	3.47	9.15	6.76	19.06

Group OC/D
(98 students in 5 schools)

DECILES	PT. I (25)	PT. II (30)	PT. III (41)	TOTAL (96)
9	8	19	18	39
8	7	17	15	33
7	6	14	13	30
6	6	13	11	28
5	4	11	9	26
4	2	10	7	23
3	2	9	6	19
2	1	7	5	15
1	0	4	1	11
Means	4.04	11.52	9.82	25.06

Group 1A
(44 students in 3 colleges)

DECILES	PT. I (25)	PT. II (30)	PT. III (41)	TOTAL (96)
9	9	14	17	33
8	8	12	15	28
7	6	10	11	25
6	4	10	9	23
5	4	9	8	21
4	3	7	7	18
3	2	6	5	16
2	1	3	2	13
1	0	1	1	10
Means	4.77	8.39	8.82	21.86

Group 1B
(1808 students in 21 colleges)

DECILES	PT. I (25)	PT. II (30)	PT. III (41)	TOTAL (96)
9	10	19	19	42
8	7	17	16	36
7	6	14	13	31
6	5	13	11	28
5	4	11	10	25
4	3	10	8	22
3	2	8	6	19
2	1	6	4	18
1	0	4	2	12
Means	4.40	11.58	10.26	26.38

Group 2
(196 students in 11 colleges)

DECILES	PT. I (25)	PT. II (30)	PT. III (41)	TOTAL (96)
9	13	22	25	58
8	11	21	21	47
7	10	18	17	40
6	7	17	14	36
5	6	14	12	32
4	4	12	10	27
3	4	11	7	24
2	2	9	5	20
1	0	6	3	14
Means	6.71	14.29	13.11	33.86

FRENCH AURAL COMPREHENSION NORMS
UPPER LEVEL FORM A

Group 10D
(72 students in 3 schools)

DECILES	PT.I (25)	PT.II (40)	PT.III (15)	TOTAL (80)
9	10	23	6	35
8	8	20	4	30
7	7	18	4	27
6	6	16	3	25
5	4	14	3	22
4	2	12	2	18
3	2	11	2	15
2	1	7	0	11
1	0	6	0	9
Means	4.93	14.36	2.88	21.91

Group 10F
(30 students in 3 schools)

DECILES	PT.I (25)	PT.II (40)	PT.III (15)	TOTAL (80)
9	16	25	6	44
8	13	23	6	37
7	12	20	3	33
6	8	19	3	29
5	7	17	2	28
4	6	15	2	26
3	6	14	2	23
2	4	11	0	20
1	2	7	0	13
Means	8.47	17.30	2.87	28.60

Group 11
(777 students in 14 colleges)

DECILES	PT.I (25)	PT.II (40)	PT.III (15)	TOTAL (80)
9	16	29	8	49
8	14	24	6	41
7	12	21	4	35
6	10	19	3	31
5	8	18	2	27
4	6	15	2	24
3	6	13	0	21
2	4	11	0	18
1	2	7	0	13
Means	8.70	17.80	3.14	29.32

Group 12
(141 students in 7 colleges)

DECILES	PT.I (25)	PT.II (40)	PT.III (15)	TOTAL (80)
9	22	36	11	65
8	20	33	9	59
7	19	30	8	53
6	16	25	6	48
5	14	23	6	43
4	12	20	4	36
3	10	18	3	32
2	8	16	2	26
1	5	12	0	21
Means	13.98	23.62	5.65	42.97

FRENCH AURAL COMPREHENSION NORMS
LOWER LEVEL FORM B

Group 1B
(71 students in 1 college)

DECILES	PT. I (25)	PT. II (30)	PT. III (35)	TOTAL (90)
9	12	21	16	43
8	10	18	11	38
7	8	15	11	31
6	7	14	9	29
5	6	14	8	27
4	5	12	7	25
3	4	9	4	21
2	2	8	3	18
1	1	7	1	15
Means	6.62	13.03	8.02	27.46

Group 2
(54 students in 1 college)

DECILES	PT. I (25)	PT. II (30)	PT. III (35)	TOTAL (90)
9	12	22	19	43
8	10	18	16	39
7	8	17	14	35
6	8	15	11	31
5	6	13	9	27
4	5	11	8	25
3	4	10	7	22
2	2	9	5	19
1	2	6	3	17
Means	6.85	13.54	10.46	30.85

FRENCH AURAL COMPREHENSION NORMS
UPPER LEVEL FORM B

Group 11
(344 students in 5 colleges)

DECILES	PT. I (25)	PT. II (40)	PT. III (15)	TOTAL (80)
9	10	20	9	34
8	8	17	8	29
7	7	14	6	25
6	6	12	5	22
5	5	11	4	19
4	4	9	3	17
3	3	7	2	15
2	2	5	0	12
1	0	3	0	9
Means	5.17	11.30	4.38	20.62

Group 12
(104 students in 3 colleges)

DECILES	PT. I (25)	PT. II (40)	PT. III (15)	TOTAL (80)
9	15	28	11	53
8	14	24	10	44
7	12	22	9	39
6	11	18	8	35
5	9	16	7	31
4	8	13	6	28
3	7	11	4	24
2	6	8	2	20
1	4	6	0	15
Means	9.64	16.56	6.56	32.71

GERMAN AURAL COMPREHENSION NORMS
LOWER LEVEL FORM A

Group OB
(160 students in 6 schools)

DECILES	PT. I (25)	PT. II (25)	PT. III (36)	TOTAL (86)
9	14	13	19	43
8	9	8	13	28
7	7	6	9	20
6	5	4	8	15
5	4	4	5	12
4	2	2	3	10
3	1	1	2	8
2	0	0	1	6
1	0	0	0	3
Means	5.27	4.95	7.42	17.47

Group OC
(32 students in 1 school)

DECILES	PT. I (25)	PT. II (25)	PT. III (36)	TOTAL (86)
9	14	12	11	29
8	8	7	9	20
7	6	5	7	14
6	2	4	3	11
5	1	2	3	8
4	0	1	2	5
3	0	1	2	4
2	0	0	2	3
1	0	0	1	1
Means	3.88	4.12	5.32	13.09

Group OD
(32 students in 4 schools)

DECILES	PT. I (25)	PT. II (25)	PT. III (36)	TOTAL (86)
9	16	14	26	47
8	14	12	15	40
7	13	11	14	35
6	12	10	10	34
5	10	7	8	26
4	7	6	7	20
3	6	5	3	14
2	4	4	1	10
1	2	1	0	7
Means	9.41	8.12	10.00	27.25

Group 1A
(300 students in 3 colleges)

DECILES	PT. I (25)	PT. II (25)	PT. III (36)	TOTAL (86)
9	12	12	17	35
8	8	8	10	23
7	6	6	8	19
6	4	5	7	16
5	4	4	5	13
4	2	2	4	10
3	1	2	2	8
2	0	1	1	6
1	0	0	0	2
Means	4.74	4.68	6.70	15.82

Group 1B
(470 students in 13 colleges)

DECILES	PT. I (25)	PT. II (25)	PT. III (36)	TOTAL (86)
9	14	16	22	47
8	12	12	16	37
7	9	10	13	30
6	7	8	10	23
5	6	7	8	19
4	4	6	6	16
3	3	4	4	12
2	2	2	2	9
1	0	0	1	6
Means	6.78	7.26	9.64	23.41

Group 2
(171 students in 6 colleges)

DECILES	PT. I (25)	PT. II (25)	PT. III (36)	TOTAL (86)
9	15	16	21	49
8	13	13	18	39
7	10	11	15	33
6	8	8	12	29
5	8	7	9	23
4	7	6	8	19
3	6	4	4	15
2	2	2	2	11
1	1	1	1	8
Means	8.25	8.10	10.62	26.53

GERMAN AURAL COMPREHENSION NORMS
UPPER LEVEL FORM A

DECILES	PT. I (25)	PT. II (40)	PT. III (15)	TOTAL (80)
(22 students in 2 schools)				
9	13	27	9	46
8	12	23	8	41
7	10	21	8	37
6	9	19	6	35
5	7	13	4	23
4	6	7	3	17
3	2	6	2	11
2	0	3	0	5
1	0	1	0	3
Means	6.86	14.22	5.09	25.81
(84 students in 3 colleges)				
9	13	30	10	53
8	10	23	9	40
7	7	19	8	29
6	6	16	6	26
5	6	13	6	22
4	4	12	4	19
3	2	9	4	16
2	2	7	3	14
1	0	4	2	12
Means	5.90	15.22	5.95	26.89

RUSSIAN AURAL COMPREHENSION NORMS
LOWER LEVEL FORM A

(157 students in 11 colleges)

DECILES	PT.I (25)	PT.II (25)	PT.III (36)	TOTAL (86)
9	21	22	32	72
8	18	19	28	65
7	13	17	23	49
6	10	15	19	42
5	8	13	17	37
4	7	11	15	32
3	6	10	12	26
2	4	8	9	22
1	1	6	5	17
Means	9.83	13.32	17.90	40.78

RUSSIAN AURAL COMPREHENSION NORMS
UPPER LEVEL FORM A

(60 students in 7 colleges)

DECILES	PT.I (25)	PT.II (34)	PT.III (15)	TOTAL (74)
9	18	31	15	62
8	16	28	12	55
7	14	26	10	50
6	12	25	10	48
5	12	23	9	41
4	10	20	8	37
3	8	16	4	32
2	6	14	3	27
1	4	11	2	19
Means	11.17	21.68	7.97	41.05

SPANISH AURAL COMPREHENSION NORMS
LOWER LEVEL FORM A

DECILES	PT. I (25)	PT. II (25)	PT. III (38)	TOTAL (88)	DECILES	PT. I (25)	PT. II (25)	PT. III (38)	TOTAL (88)
Group OA (355 students in 5 schools)					*Group OB* (743 students in 11 schools)				
9	8	12	17	31	9	8	12	17	31
8	6	10	13	26	8	7	10	14	26
7	5	8	11	23	7	6	8	11	22
6	4	7	10	20	6	4	7	9	20
5	3	6	8	17	5	3	6	7	17
4	2	5	6	15	4	2	5	5	14
3	1	4	5	13	3	1	4	3	12
2	0	2	3	10	2	0	2	0	9
1	0	1	1	8	1	0	1	0	6
Means	3.61	6.53	8.76	18.70	Means	3.75	6.51	8.14	18.18
Group OC (158 students in 3 schools)					*Group OD* (365 students in 10 schools)				
9	10	14	21	39	9	14	16	23	47
8	8	12	18	34	8	10	14	18	40
7	6	10	15	30	7	8	12	15	35
6	6	10	13	28	6	7	12	13	30
5	4	8	11	24	5	6	10	11	26
4	4	7	10	22	4	4	8	10	24
3	2	6	8	19	3	4	8	8	20
2	2	6	5	15	2	2	6	6	17
1	0	3	3	11	1	0	4	3	12
Means	4.87	8.40	11.86	24.94	Means	6.41	10.41	12.00	28.34
Group OE (27 students in 2 schools)					*Group OF* (36 students in 3 schools)				
9	17	18	25	51	9	20	22	29	69
8	12	16	19	43	8	19	20	27	65
7	10	14	17	40	7	16	19	25	59
6	7	13	13	31	6	15	18	23	55
5	6	10	12	29	5	14	16	22	52
4	4	10	11	25	4	13	16	17	47
3	3	8	10	22	3	12	16	16	44
2	2	7	5	19	2	11	13	14	39
1	0	5	4	13	1	10	11	9	34
Means	7.33	11.67	13.68	32.30	Means	14.80	17.08	20.22	51.94

SPANISH AURAL COMPREHENSION NORMS
LOWER LEVEL FORM A

DECILES	PT. I (25)	PT. II (25)	PT. III (38)	TOTAL (88)
Group 1A (31 students in 2 colleges)				
9	8	13	17	33
8	7	10	16	27
7	6	8	15	25
6	4	7	12	23
5	4	6	11	21
4	4	6	10	20
3	2	4	8	18
2	1	4	6	16
1	0	4	5	13
Means	4.39	7.22	11.46	22.90
Group 2 (582 students in 16 colleges)				
9	16	20	27	59
8	12	18	23	50
7	10	16	20	43
6	8	14	17	38
5	7	12	16	34
4	6	11	13	29
3	4	10	11	26
2	2	8	9	22
1	1	6	5	17
Means	7.67	12.86	15.68	35.98

DECILES	PT. I (25)	PT. II (25)	PT. III (38)	TOTAL (88)
Group 1B (1412 students in 22 colleges)				
9	12	17	23	47
8	9	14	19	39
7	7	13	17	34
6	6	12	15	31
5	5	10	13	28
4	4	9	11	25
3	2	8	10	22
2	2	6	7	19
1	0	4	4	14
Means	5.64	10.48	13.42	29.38

SPANISH AURAL COMPREHENSION NORMS
UPPER LEVEL FORM A

Group 10C
(61 students in 3 schools)

DECILES	PT. I (25)	PT. II (45)	PT. III (15)	TOTAL (85)
9	12	24	6	36
8	10	22	4	33
7	9	20	3	31
6	7	19	2	28
5	6	16	2	26
4	6	15	0	23
3	4	13	0	19
2	2	9	0	16
1	1	5	0	11
Means	7.06	16.10	2.31	25.42

Group 10E
(63 students in 2 schools)

DECILES	PT. I (25)	PT. II (45)	PT. III (15)	TOTAL (85)
9	10	25	6	37
8	9	23	4	31
7	8	21	4	29
6	7	19	3	27
5	6	16	3	25
4	5	15	2	24
3	4	14	2	21
2	2	11	0	20
1	0	9	0	16
Means	6.28	17.16	3.06	26.50

Group 10G/H
(47 students in 3 schools)

DECILES	PT. I (25)	PT. II (45)	PT. III (15)	TOTAL (85)
9	18	35	10	58
8	16	32	8	53
7	13	28	6	46
6	10	26	4	40
5	9	25	4	37
4	8	21	3	33
3	5	19	2	29
2	3	15	1	25
1	2	14	0	20
Means	9.83	24.54	4.66	38.78

Group 10D
(418 students in 6 schools)

DECILES	PT. I (25)	PT. II (45)	PT. III (15)	TOTAL (85)
9	13	25	8	38
8	10	21	6	33
7	8	18	4	30
6	8	16	3	26
5	7	14	3	23
4	6	12	2	20
3	4	9	2	17
2	4	7	0	14
1	2	3	0	10
Means	7.14	14.02	3.13	24.10

Group 10F
(140 students in 8 schools)

DECILES	PT. I (25)	PT. II (45)	PT. III (15)	TOTAL (85)
9	16	28	8	45
8	13	24	6	41
7	12	21	6	36
6	10	19	4	33
5	8	17	4	30
4	8	15	3	27
3	7	14	3	25
2	6	11	2	22
1	4	8	0	17
Means	9.48	17.46	4.14	30.94

SPANISH AURAL COMPREHENSION NORMS
UPPER LEVEL FORM A

DECILES	PT. I (25)	PT. II (45)	PT. III (15)	TOTAL (85)
		Group 11		
		(1044 students in 19 colleges)		
9	**16**	**33**	**9**	**53**
8	**13**	**28**	**8**	**46**
7	**12**	**26**	**6**	**40**
6	**10**	**24**	**4**	**37**
5	**8**	**21**	**4**	**34**
4	**8**	**20**	**3**	**31**
3	**6**	**17**	**2**	**28**
2	**5**	**15**	**2**	**24**
1	**2**	**11**	**0**	**20**
Means	**9.37**	**21.68**	**4.34**	**35.22**
		Group 12		
		(302 students in 10 colleges)		
9	**20**	**35**	**9**	**61**
8	**16**	**31**	**7**	**52**
7	**14**	**28**	**6**	**47**
6	**13**	**27**	**4**	**42**
5	**12**	**25**	**4**	**38**
4	**10**	**21**	**3**	**35**
3	**8**	**20**	**2**	**31**
2	**6**	**17**	**2**	**27**
1	**4**	**14**	**0**	**22**
Means	**11.48**	**24.29**	**4.36**	**40.06**

SPANISH AURAL COMPREHENSION NORMS
LOWER LEVEL FORM B

Group 1B
(141 students in 4 colleges)

DECILES	PT.I (25)	PT.II (25)	PT.III (38)	TOTAL (88)
9	13	16	23	47
8	10	14	20	43
7	9	12	18	38
6	8	10	16	33
5	6	9	14	30
4	5	8	13	27
3	4	7	11	23
2	2	6	9	18
1	0	2	6	14
Means	6.69	9.40	14.44	30.33

Group 2
(165 students in 3 colleges)

DECILES	PT.I (25)	PT.II (25)	PT.III (38)	TOTAL (88)
9	18	19	28	63
8	15	18	24	53
7	13	16	22	48
6	12	14	19	43
5	10	12	16	36
4	8	10	14	32
3	6	9	10	28
2	4	7	9	23
1	3	4	5	18
Means	10.09	12.10	16.48	38.66

SPANISH AURAL COMPREHENSION NORMS
UPPER LEVEL FORM B

Group 11
(303 students in 3 colleges)

DECILES	PT.I (25)	PT.II (44)	PT.III (15)	TOTAL (84)
9	15	16	10	38
8	13	13	9	32
7	12	11	8	28
6	10	9	8	25
5	10	7	7	22
4	8	5	5	20
3	7	4	4	17
2	6	2	2	14
1	3	1	0	10
Means	9.29	8.10	6.04	23.86

Group 12
(127 students in 4 colleges)

DECILES	PT.I (25)	PT.II (44)	PT.III (15)	TOTAL (84)
9	19	26	14	56
8	18	22	12	47
7	16	19	10	44
6	15	17	9	40
5	14	15	9	37
4	13	12	8	33
3	12	10	6	28
2	10	8	4	24
1	7	4	2	18
Means	13.53	14.72	8.16	36.18

RUSSIAN READING-VOCABULARY-GRAMMAR TEST NORMS

(173 students in 7 colleges)

DECILES	READ. (45)	VOC. (40)	GRAM. (40)	TOTAL (125)
9	39	27	30	92
8	31	21	27	80
7	26	17	23	65
6	21	15	19	55
5	17	13	17	45
4	15	9	15	40
3	12	8	11	30
2	9	6	10	25
1	5	4	7	18
Means	19.82	13.94	17.74	51.00

Appendix B

Charts of Co-operative and Aural Test Data

The following charts contain all Co-operative and aural scores obtained from the colleges and secondary schools which collaborated specially with the Investigation, and whose programs are discussed in Chapters V and VI of the present volume. In cases where scores of both experimental and conventional groups appear in the same chart (e.g. extensively for College J), experimental groups are distinguished from their conventional counterparts by the symbol X- preceding their group number. Thus for example, in College J French, Experimental Group la is designated as X-la, as contrasted with Conventional Group la, etc.

COLLEGE A	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH
SPANISH	1	5	84	JAN. 1947	75	FORM P	1 YR.	LOWER B		MAY 1947	150	FORM Q	1 YR.	LOWER A	1B
						R 43.4	13	I 3.5	-			R 60.9	77	I 4.0	4
						V 43.7	11	II 4.7	-			V 59.4	64	II 8.8	4
						G 45.6	25	III 7.8	-			G 54.6	61	III 14.0	6
						T 43.8	13	T 16.1	-			T 59.2	68	T 26.7	4
	2	4	58	OCT. 1946	150*	FORM P	1 YR.	LOWER B	1B	MAY 1947	270	FORM Q	2 YRS.	LOWER A	2
						R 59.4	71	I 8.4	6			R 66.2	51	I 8.6	7
						V 52.6	39	II 9.9	6			V 63.4	35	II 13.8	6
						G 56.0	65	III 13.6	5			G 66.2	64	III 18.6	7
						T 56.7♦	60	T 32.0♦	6			T 66.9♦	53	T 41.0♦	7
	2.1#		35			R 58.9	71	I 8.4	6			R 64.6	47	I 7.8	6
						V 51.8	35	II 8.9	5			V 60.9	28	II 13.2	6
						G 53.7	57	III 12.9	4			G 63.8	56	III 18.4	7
						T 55.6	56	T 30.2	5			T 64.6	44	T 39.4	7
	3	2	16	JAN. 1947	300	FORM P	3 YRS.	LOWER B	2	MAY 1947	330	FORM Q	3 YRS.	LOWER A	2
						R 74.2	52	I 13.6	8						
						V 71.1	32	II 15.9	7						
						G 72.7	74	III 20.0	7						
						T 75.0	55	T 49.5	7						
	3.1⊕		7			R 75.0	57	I 11.4	5			R 70.9	40	I 11.1	8
						V 66.3	16	II 14.9	6			V 74.9	47	II 17.6	8
						G 71.1	67	III 16.6	5			G 70.3	64	III 21.9	8
						T 73.0	47	T 42.9	6			T 74.1	51	T 50.6	8
FRENCH	1A	5	51	JAN. 1947	75	FORM Q	1 YR.	LOWER B		MAY 1947	150	FORM R	1 YR.	LOWER A	1B
						R 45.5	19	I 4.2	-			R 59.1	66	I 5.1	6
						V 57.2	39	II 10.7	-			V 60.0	52	II 13.0	7
						G 52.9	45	III 6.1	-			G 60.8	76	III 10.2	5
						T 51.9	32	T 21.0	-			T 61.0	67	T 28.4	7
	1B	6	16	JAN. 1947	90	FORM Q	1 YR.	LOWER B	1B	MAY 1947	180	FORM R	1 YR.	LOWER A	1B
						R 42.6	13	I 2.9	3			R 65.4	84	I 3.3	4
						V 49.8	16	II 8.1	2			V 60.3	53	II 9.1	4
						G 52.0	40	III 6.3	4			G 59.8	73	III 10.0	5
						T 47.9	19	T 17.3	2			T 63.1	74	T 22.4	4
RUSSIAN	1	5	12							MAY 1947	150	INV.	TENTH	LOWER A	
												R 13.7	3	I 4.7	3
												V 5.3	1	II 10.8	4
												G 9.8	3	III 16.8	5
												T 28.8	2	T 32.3	5
	2	4	6	JAN. 1947	210	INV.	TENTH	LOWER A		MAY 1947	270	INV.	TENTH	LOWER A	
						R 13.0	3	I 2.3	2			R 16.7	5	I 2.7	2
						V 9.8	5	II 8.3	2			V 11.7	5	II 4.7	1
						G 8.8	2	III 6.0	2			G 10.7	3	III 10.7	3
						T 31.7	3	T 16.7	1			T 39.0¶	4	T 18.0¶	2

*Or the high-school equivalent.

♦Excluding Subgroup 2.1, total means are, for 150 hrs.: Co-operative 57.9, aural 32.5; for 270 hrs.: Co-operative 70.0, aural 42.9.

#Twenty-one members of this subgroup had taken the lower-level aural (Form A) in May 1946 at 150 hrs., with the following mean scores: I 7.2, II 12.9, III 13.1, T 33.1.

⊕This subgroup continued through 30 more hrs. and was tested again in May 1947.

¶Only three of the group completed the 270 hrs. Total mean scores for these three at 210 hrs. had been: reading, 30.3; aural, 17.3.

COLLEGE B		HRS. PER			TOTAL	CO-OPERATIVE		AURAL			TOTAL	CO-OPERATIVE		AURAL	
Conventional	GROUP	WK.	N	DATE	HRS.	MEAN	%ILE	MEAN	TENTH	DATE	HRS.	MEAN	%ILE	MEAN	TENTH
FRENCH	2A	3	50							OCT. 1945	145	FORM R R 64.7 V 66.9 G 63.9 T 66.6	2 YRS. 40 28 61 42	UPPER A I 6.4 II 16.4 III 3.1 T 25.9	11 4 5 6 4
	2B	3	76	OCT. 1945	105*	FORM R R 62.8 V 66.6 G 66.0 T 66.7	1 YR. 78 77 89 85	UPPER A I 5.6 II 15.8 III 2.3 T 23.7	11 4 5 5 4	APR. 1946	160	FORM R R 71.8 V 71.6 G 68.0 T 72.7	2 YRS. 67 47 76 65	UPPER B I 4.0 II 11.2 III 4.0 T 19.2	11 4 5 5 5
	2C	3	33	FEB. 1946	95*	FORM R R 58.3 V 66.3 G 62.7 T 63.8	1 YR. 63 76 83 76	LOWER B I 8.0 II 14.7 III 7.6 T 30.3	1B 7 7 5 7	JUNE 1946	140	FORM R R 68.9 V 71.0 G 66.9 T 71.0	2 YRS. 61 45 72 59	LOWER A I 8.4 II 14.8 III 12.7 T 35.9	2 7 6 6 6
	3A	3	17	OCT. 1945	H.S. +15	FORM R R 79.9 V 76.8 G 67.8 T 76.9	2 YRS. 89 68 75 80	UPPER A I 13.9 II 23.1 III 2.7 T 39.7	12 5 5 3 5	APR. 1946	H.S. +70	FORM R R 85.6 V 79.2 G 70.8 T 82.1	3 YRS. 85 40 73 72	UPPER B I 8.6 II 18.4 III 7.3 T 34.3	12 5 6 5 6
	3B	3	11							APR. 1946	235	FORM R R 81.9 V 79.4 G 74.5 T 81.5	3 YRS. 75 40 84 70	UPPER B I 7.7 II 17.1 III 4.3 T 29.1	12 4 5 3 4
SPANISH	2A	3	63							NOV. 1945	145	FORM Q R 65.8 V 65.0 G 67.0 T 67.5	2 YRS. 50 42 68 55	UPPER A I 9.0 II 22.1 III 3.4 T 34.5	11 5 6 4 5
	2B	3	127	NOV. 1945	105*	FORM Q R 59.9 V 59.3 G 61.9 T 61.5	1 YR. 74 65 85 77	UPPER A I 5.5 II 18.8 III 3.7 T 28.0	11 4 4 5 4	APR. 1946	160	FORM Q R 66.8 V 66.3 G 66.8 T 68.4	2 YRS. 54 48 67 59	UPPER A I 8.1 II 24.9 III 4.8 T 37.8	11 5 6 6 6
	2C	3	16	FEB. 1946	95*	FORM Q R 60.0 V 58.2 G 59.5 T 60,2	1 YR. 74 61 77 73	LOWER B I 8.6 II 11.0 III 18.7 T 38,3	1B 7 6 7 7	JUNE 1946	140	FORM Q R 66.6 V 67.0 G 63.3 T 67.3	2 YRS. 54 50 53 54	LOWER A I 9.9 II 14.4 III 20.4 T 44.7	2 7 6 7 8

*Or the high-school equivalent.

COLLEGE D	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH	HRS. PER WK.	N	DATE	TOTAL HRS.	AURAL MEAN	TENTH
FRENCH	1	6	77	JAN. 1946	90	FORM Q R 48.0 V 60.0 G 59.8 T 56.3	1 YR. 26 52 73 47	LOWER A I 3.2 II 8.5 III 9.2 T 20.9	1B 4 3 5 4	6	60	MAY 1946	180	LOWER A I 3.6 II 11.5 III 12.2 T 27.3	1B 5 5 7 5
	2	6	33	JAN. 1946	280*	FORM Q R 72.8 V 77.2 G 67.7 T 74.7	2 YRS. 71 70 75 73			4	33	MAY 1946	340	LOWER A I 3.5 II 13.9 III 12.6 T 30.0	2 4 5 5 5
GERMAN	1	6	34	JAN. 1946	90	FORM P R 51.9 V 54.9 G 52.9 T 53.2	1 YR. 34 44 51 40	LOWER A I 4.8 II 6.1 III 6.5 T 17.4	1B 5 4 4 5	6	30	MAY 1946	180	LOWER A I 10.4 II 9.9 III 13.0 T 33.3	1B 8 7 7 8
	2	6	12	JAN. 1946	280*	FORM P R 69.1 V 65.8 G 54.1 T 64.0	2 YRS. 58 43 19 37			4	12	MAY 1946	340	LOWER A I 5.8 II 7.9 III 7.0 T 20.8	2 3 6 4 5
SPANISH	1	6	40	JAN. 1946	90	FORM Q R 52.3 V 51.2 G 54.3 T 52.7	1 YR. 42 32 57 44	LOWER A I 3.2 II 8.3 III 10.4 T 22.0	1B 4 3 3 3	6	20	MAY 1946	180	LOWER A I 7.2 II 10.8 III 15.8 T 33.8	1B 7 6 7 7
	2	6	47	JAN. 1946	280*	FORM Q R 59.7 V 61.3 G 73.9 T 66.2	2 YRS. 26 28 89 49			4	47	MAY 1946	340	LOWER A I 8.5 II 15.3 III 16.0 T 39.8	2 6 7 6 7

*Or the high-school equivalent.

COLLEGE C June 1946	GROUP	HRS. PER WK.	N	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH	Further Aural Notes
Spanish	1	3	56	90	FORM P R 60.1 V 59.1 G 66.7 T 63.4	1 YR. 74 64 94 82	LOWER B I 5.7 II 8.5 III 13.6 T 27.8*	1B 5 5 5 5	1) Twenty-two of these (*27.2) took Aural Lower A at 135 hrs. in Jan. 1947. Total mean: 37.1 (6th tenth). 2) Twelve others (*23.7) took Aural Upper A at 180 hrs. in Jan. 1947. Total mean: 30.5 (4th tenth).
French	2	3	53	135	FORM Q R 63.9 V 69.7 G 61.2 T 66.5	2 YRS. 38 38 49 41	LOWER B I 6.9 II 13.8 III 10.5 T 31.1*	2 6 5 6 6	1) Nineteen of these (*32.4) took Aural Upper A at 180 hrs. in Sept. 1946. Total mean: 44.2 (9th tenth). 2) Fourteen others (*27.4) took Aural Upper A at 180 hrs. in Jan. 1947. Total mean: 26.3 (4th tenth).
Spanish	2	3	35	135	FORM P R 66.5 V 64.3 G 69.2 T 68.6	2 YRS. 53 40 76 59	LOWER B I 9.8 II 12.8 III 17.6 T 40.2*	2 5 5 6 6	1) Eleven of these (*50.5) had taken Aural Lower A at 90 hrs. in Feb. 1946. Total mean: 34.3 (7th tenth). 2) Four of the above eleven (*51.5) took Aural Upper A at 180 hrs. in Jan. 1947. Total mean: 32.3 (5th tenth). 3) Eleven others (*38.0) took Aural Upper A at 180 hrs. in Jan. 1947. Total mean: 37.9 (6th tenth).
French	3	3	23	180	FORM Q R 65.5 V 70.1 G 62.0 T 67.0	2 YRS. 43 40 52 43	UPPER B I 4.1 II 9.5 III 3.3 T 16.9*	11 4 5 4 4	1) Thirteen of these (*19.6) had taken Aural Lower A at 135 hrs. in Feb. 1946. Total mean: 33.6 (5th tenth).
Spanish	3	3	25	180	FORM P R 69.4 V 67.9 G 69.7 T 70.8	2 YRS. 65 54 78 68	UPPER B I 11.1 II 7.9 III 5.1 T 24.1*	11 5 6 4 6	1) Fifteen of these (*25.3) had taken Aural Lower A at 135 hrs. in Feb. 1946. Total mean: 41.8 (sic) (7th tenth). 2) Five of the above fifteen (*25.8) took Aural Upper A at 225 hrs. in Jan. 1947. Total mean: 41.8 (sic) (5th tenth). 3) One other (*22.0) took Aural Upper A at 225 hrs. in Jan. 1947. Total score: 58.0 (8th tenth).
French	4	3	27	225 OR MORE	FORM Q R 82.1 V 79.6 G 71.1 T 80.6	3 YRS. 75 41 74 66	UPPER B I 9.6 II 17.1 III 6.7 T 33.3*	12 6 5 5 6	1) Five of these (*41.2) took Aural Upper A after 45 more hrs. in Sept. 1946. Total mean: 59.0 (8th tenth). 2) Five others (*47.8) took Aural Upper A after 45 more hrs. in Jan. 1947. Total mean: 58.4 (8th tenth).
Spanish	4	3	30	225 OR MORE	FORM P R 73.6 V 77.4 G 75.3 T 78.0	3 YRS. 50 56 81 67	UPPER B I 15.1 II 16.9 III 8.8 T 40.9*	12 5 6 5 6	1) Seven of these (*39.1) took Aural Upper A after 45 more hrs. in Jan. 1947. Total mean: 46.3 (6th tenth).

COLLEGE E	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
French	1	4½	107	MAY 1946	135	FORM R R 57.1 V 56.5 G 56.0 T 57.7	1 YR. 59 35 57 55	LOWER A I 3.8 II 10.2 III 7.9 T 21.9	1B 5 4 4 4
Spanish	1	4½	129	MAY 1946	135	FORM Q R 61.8 V 57.5 G 54.2 T 58.7	1 YR. 80 60 57 68	LOWER A I 5.9 II 11.3 III 13.7 T 30.9	1B 6 6 6 5
French	2	4½	47	MAY 1946	225*	FORM R R 80.0 V 77.2 G 71.3 T 78.9	2 YRS. 89 69 86 86	UPPER A I 9.5 II 17.1 III 4.0 T 30.5	11 6 5 7 6
Spanish	2	4½	74	MAY 1946	225*	FORM Q R 70.5 V 75.8 G 69.5 T 74.2	2 YRS. 67 82 79 79	UPPER A I 9.9 II 21.9 III 4.9 T 36.7	11 6 6 6 6

*Or high-school equivalent of first 90.

COLLEGE F	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
German	1	3	128	JUNE 1947	50	FORM* R 46.9 V 53.2 G 51.2 T 50.2	 - - - -	LOWER A I 3.8 II 4.4 III 5.0 T 13.2	1A 5 5 5 5

*N, O, or P.

COLLEGE G	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
FRENCH	1A*	4	31	DEC. 1944	35	FORM R R 58.0 V 68.7 G 50.4 T 59.4	MAR. 1945	70	FORM R R 79.2 V 78.2 G 60.1 T 75.1	JUNE 1945	105	FORM R R 84.1 V 82.4 G 57.6 T 77.3	1 YR. 99 99 65 99	LOWER A I 6.4 II 16.7 III 11.6 T 34.7	1B 7 8 7 8
	1A.1♦		8		35+	R 59.9 V 65.2 G 56.4 T 60.9		70+	R 75.5 V 76.7 G 65.2 T 74.8		105+	R 81.5 V 80.4 G 60.4 T 76.5	99 98 73 98	I 6.1 II 17.6 III 6.8 T 30.5	7 8 3 7
	1B	4	13	DEC. 1945	35	FORM R R 57.3 V 64.8 G 49.1 T 57.7	MAR. 1946	70	FORM R R 77.6 V 79.2 G 55.3 T 72.8	JUNE 1946	105	FORM R R 84.0 V 84.1 G 54.3 T 76.5	1 YR. 99 99 49 98	LOWER A I 8.9 II 18.9 III 10.9 T 38.6	1B 9 9 5 8
	1C.1#	12	13	JULY 22 1946	35	FORM R R 55.8 V 65.7 G 53.3 T 59.2	AUG. 7 1946	70	FORM R R 73.5 V 76.1 G 55.5 T 70.3	AUG. 29 1946	105	FORM R R 82.5 V 81.0 G 58.2 T 76.5	1 YR. 99 99 65 98	LOWER A I 6.7 II 13.2 III 15.0 T 34.9	1B 8 7 8 8
	1C.2♦	12	15		35+	R 59.2 V 68.9 G 55.0 T 62.2		70+	R 79.0 V 74.7 G 59.4 T 73.2		105+	R 85.0 V 80.5 G 59.3 T 77.6	99 98 69 99	I 7.3 II 13.3 III 13.4 T 34.1	8 7 7 8
	1D⊕	4	16				JUNE 1946	70	FORM R R 81.5 V 75.2 G 56.3 T 73.3	AUG. 1946	105	FORM R R 88.0 V 81.8 G 56.9 T 78.4	1 YR. 99 99 61 99	LOWER A I 10.2 II 16.9 III 17.3 T 44.3	1B 9 8 8 10
SPANISH	1¶	12	20	JULY 29 1946	35	FORM O R 60.6 V 62.7 G 64.2 T 63.9	AUG. 14 1946	70	FORM O R 69.6 V 68.1 G 65.1 T 69.5	AUG. 28 1946	105	FORM O R 75.3 V 79.8 G 71.3 T 77.8	1 YR. 99 99 98 99	LOWER A I 10.1 II 14.6 III 18.8 T 43.5	1B 9 9 8 9

*Excluding subgroup 1a.1, total means are: Co-operative at 35 hrs., 58.9; at 70 hrs., 75.2; at 105 hrs., 77.6; aural, 36.1.

♦This subgroup had previous study of French.

#This subgroup had no previous study of French.

⊕This group consisted of war veterans.

¶Some of this group had previous study of Spanish.

HIGH SCHOOL A	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
FRENCH	1	4	33					JUNE 1946	100	FORM R* R 49.4 V 47.9 G 48.5 T 48.1	1 YR. 75 53 52 62	LOWER A I 4.9 II 9.5 III 5.7 T 20.1	OB 7 6 5 6
	2	4	27	OCT. 1945	110	LOWER A I 4.9 II 14.8 III 9.1 T 28.7	OB 7 8 7 9	JUNE 1946	200	FORM Q R 67.6 V 64.7 G 64.7 T 67.2	2 YRS. 84 70 74 77	LOWER B I 12.2 II 17.6 III 10.0 T 39.8	 - - - -
	3	4	9	OCT. 1945	210	UPPER A I 11.0 II 14.4 III 3.4 T 28.9	10D 9 5 6 7	JUNE 1946	300	FORM Q R 83.3 V 76.9 G 73.6 T 80.8	3 YRS. 93 73 78 86	UPPER B I 7.2 II 16.2 III 8.0 T 31.4	 - - - -
SPANISH	2	4	28					JUNE 1946	200	FORM Q R 47.3 V 46.4 G 51.8 T 48.6	2 YRS. 24 21 42 30	LOWER A I 3.8 II 8.1 III 12.5 T 24.3	OD 4 4 6 4

*Elementary Form.

COLLEGE H 1945-1946	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH
GERMAN	X-1A	8	34	FEB. 1946	120	FORM Q	1 YR.			JUNE 1946	240	FORM N	2 YRS.		
						R 49.9	27								
						V 45.0	11								
						G 51.0	43								
						T 48.4	23								
	X1A.1*		21			R 49.4	24					R 62.9	35		
						V 44.4	10					V 56.5	13		
						G 49.8	39					G 56.0	25		
						T 47.6	22					T 59.2	20		
	X-1B	8	33	JUNE 1946	120	FORM O	1 YR.								
						R 50.3	28								
						V 48.6	22								
						G 53.0	51								
						T 50.5	31								
	1A	3	33	JUNE 1946	90	FORM P	1 YR.	LOWER A	1B	AUG. 1946	180			UPPER A	11-12
						R 52.2	34								
						V 52.5	37								
						G 54.3	56								
						T 53.2	40								
	1A.1♦		8			R 55.3	45	I 8.3	7					I 8.0	8
						V 55.1	44	II 7.1	5					II 17.4	6
						G 62.3	83	III 11.4	6					III 6.0	5
						T 58.1	61	T 26.8	7					T 31.4	8
	1A.2#		10	JULY 1946				I 9.6	8					I 7.4	7
								II 8.1	6					II 17.3	6
								III 14.1	8					III 4.6	5
								T 31.8	7					T 29.3	7
	X-2A	8	20	DEC. 1945	150	FORM Q	1 YR.	LOWER A	2	FEB. 1946	240	FORM P	2 YRS.		
						R 52.2	34	I 9.1	6			R 59.4	21		
						V 51.8	32	II 8.6	7			V 62.8	32		
						G 51.4	44	III 11.8	6			G 66.1	64		
						T 51.8	36	T 29.5	7			T 64.1	37		
	X-2B	8	11							JUNE 1946	210	FORM N	2 YRS.		
												R 61.3	28		
												V 60.2	22		
												G 62.5	50		
												T 62.4	32		
	2A	3	94	DEC. 1945	100	FORM Q	1 YR.	LOWER A	1B						
								I 7.9	7						
								II 6.2	4						
								III 8.7	5						
								T 22.8	6						
	2A.1⊕		59			R 49.9	27	I 6.9	6						
						V 47.7	19	II 4.9	4						
						G 48.9	35	III 5.9	4						
						T 48.6	25	T 17.7	5						

*Only this subgroup completed 240 hrs.

♦After 90 hrs. at 3 per wk., this subgroup completed 180 hrs. at 14 per wk.

#This subgroup had the equivalent of the first 90 hrs. in high school or in an earlier German course.

⊕Only this subgroup took the Co-operative test.

COLLEGE H 1946-1947	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
French	1A	8	48	FEB. 1947	120	FORM R	1 YR.	LOWER A	1B	JUNE 1947	240	FORM Q	2 YRS.	UPPER A	11
						R 52.3	41	I 8.0	9			R 69.1	57	I 12.0	7
						V 59.1	48	II 15.1	8			V 73.5	55	II 17.8	5
						G 56.4	59	III 9.6	5			G 63.9	61	III 3.2	6
						T 56.5	49	T 32.8	8			T 70.6	58	T 33.0	7
	1A.1*		8			R 65.4	85	I 10.0	9			R 69.0	57	I 11.6	7
						V 69.1	84	II 18.1	8			V 73.0	53	II 18.9	5
						G 65.8	89	III 9.9	5			G 68.3	77	III 4.5	7
						T 68.5	89	T 38.0	8			T 72.1	63	T 35.0	7
German	1A	8	113	FEB. 1947	120	FORM Q	1 YR.	LOWER A	1B	JUNE 1947	240	FORM P	2 YRS.	UPPER A	11-12
						R 49.9	27	I 6.9	6			R 58.7	21	I 6.5	5
						V 43.2	8	II 5.3	4			V 59.5	21	II 17.4	6
						G 55.5	61	III 7.8	5			G 58.4	33	III 5.9	5
						T 49.8	29	T 20.0	5			T 59.7	22	T 29.8	8
	1A.1*		10			R 54.3	42	I 8.5	7			R 61.7	30	I 8.7	8
						V 52.1	33	II 7.1	5			V 65.0	40	II 19.4	7
						G 64.0	88	III 8.9	5			G 62.2	49	III 7.2	6
						T 57.4	57	T 24.5	7			T 64.2	57	T 35.3	8
Spanish	1A	8	70	FEB. 1947	120	FORM Q	1 YR.	LOWER A	1B	JUNE 1947	240	FORM P	2 YRS.	UPPER A	11
						R 55.3	56	I 7.8	8			R 66.1	51	I 10.9	7
						V 52.8	38	II 12.7	7			V 60.7	27	II 25.3	7
						G 50.6	42	III 14.6	6			G 64.0	56	III 3.8	6
						T 53.7	46	T 35.1	8			T 65.0	44	T 40.0	7
	1A.1*		8			R 64.4	87	I 12.1	9			R 70.8	70	I 11.0	7
						V 63.9	81	II 15.4	9			V 66.1	46	II 29.5	9
						G 60.9	82	III 20.4	9			G 71.0	82	III 4.6	7
						T 64.5	85	T 47.9	10			T 71.3	69	T 45.1	8
Russian	1A	8	19	FEB. 1947	120	INV.	TENTH	LOWER A		JUNE 1947	240	INV.	TENTH	UPPER A	
						R 5.2	1	I 5.2	3			R 20.5	6	I 6.6	3
						V 5.2	1	II 11.1	4			V 16.6	7	II 17.5	4
						G 12.9	4	III 14.9	4			G 21.1	7	III 5.4	4
						T 23.3	1	T 31.2	4			T 58.2	6	T 29.5	3
French	1B	8	41	JUNE 1947	120	FORM R	1 YR.	LOWER B	1B						
						R 57.8	62	I 8.4	7						
						V 60.9	56	II 13.0	5						
						G 58.7	68	III 7.9	5						
						T 60.8	67	T 29.4	7						
German	1B	8	106	JUNE 1947	120	FORM Q	1 YR.	LOWER A	1B						
						R 49.7	27	I 6.8	6						
						V 44.2	10	II 5.0	4						
						G 57.1	67	III 6.9	4						
						T 50.2	30	T 18.7	5						
Spanish	1B	8	61	JUNE 1947	120	FORM Q	1 YR.	LOWER B	1B						
						R 53.2	47	I 10.0	8						
						V 50.3	30	II 9.8	6						
						G 45.3	23	III 9.1	2						
						T 49.5	30	T 28.9	5						
French	2A	8	98							FEB. 1947	240	FORM Q	2 YRS.	UPPER B	11
												R 69.8	60	I 7.9	8
												V 75.3	62	II 18.3	9
												G 66.9	72	III 7.3	8
												T 72.9	66	T 33.6	9
	2A.1†		30									R 73.5	73	I 9.6	9
												V 77.8	72	II 19.1	9
												G 70.5	84	III 6.4	7
												T 76.5	79	T 35.0	9

*Total mean scores made by these subgroups at placement in October, 1946, were:
French: Co-op. R 61.1, Aural Upper A 20.5
German: Co-op. Q 47.2, Aural Upper A 17.0
Spanish: Co-op. Q56.5, Aural Upper A 26.5

COLLEGE H		HRS. PER			TOTAL	CO-OPERATIVE		AURAL			TOTAL	CO-OPERATIVE		AURAL	
1946-1947	GROUP	WK.	N	DATE	HRS.	MEAN	%ILE	MEAN	TENTH	DATE	HRS.	MEAN	%ILE	MEAN	TENTH
German	2A	8	98							FEB. 1947	240	FORM P R 62.3 V 62.2 G 61.6 T 63.1	2 YRS. 32 29 47 34	UPPER A I 5.5 II 18.7 III 6.4 T 30.6	11-12 5 7 6 8
	2A.1♦		18									R 68.0 V 69.2 G 71.3 T 71.1	54 57 81 66	I 8.0 II 28.1 III 7.8 T 43.9	8 9 7 9
Spanish	2A	8	105	♦Total mean scores made by these subgroups at placement in October, 1946, were: French: Co-op. R 67.5, Aural Upper A 30.6 German: Co-op. Q 62.0, Aural Upper A 31.5 Spanish: Co-op. Q65.2, Aural Upper A 31.8						FEB. 1947	240	FORM P R 71.2 V 68.8 G 70.5 T 72.3	2 YRS. 71 58 79 72	UPPER B I 11.7 II 9.1 III 5.4 T 26.2	11 7 6 4 7
	2A.1♦		24									R 77.0 V 79.8 G 76.5 T 80.3	88 91 93 93	I 13.8 II 11.0 III 5.8 T 30.6	9 7 5 8
French	2B	8	16							JUNE 1947	240	FORM Q R 77.6 V 79.1 G 71.0 T 78.7	2 YRS. 84 76 86 85	UPPER A I 13.7 II 20.8 III 5.0 T 39.5	11 8 7 8 8
	2B.1#		9									R 82.6 V 80.9 G 72.1 T 81.6	93 82 89 90	I 15.1 II 23.4 III 5.9 T 44.4	8 8 8 9
German	2B	8	24							JUNE 1947	240	FORM P R 62.7 V 63.5 G 60.8 T 63.5	2 YRS. 34 34 44 35	UPPER A I 6.3 II 16.3 III 6.0 T 28.6	11-12 5 6 5 7
	2B.1#		11									R 63.9 V 62.5 G 62.1 T 64.2	38 30 48 38	I 6.6 II 18.4 III 6.8 T 31.8	7 7 6 8
Spanish	2B	8	10	#Total mean scores made by these subgroups at placement in February, 1947, were: French: Co-op. Q 73.9, Aural Upper B 16.0 German: Co-op. Q 49.8, Aural Upper A 17.6 Spanish: Co-op. P 62.4, Aural Upper B 16.0						JUNE 1947	240	FORM P R 70.8 V 65.1 G 69.2 T 70.4	2 YRS. 70 42 76 67	UPPER A I 11.7 II 24.7 III 2.4 T 38.8	11 7 7 3 7
	2B.1#		7									R 72.6 V 67.4 G 72.1 T 73.1	76 52 84 76	I 12.1 II 23.7 III 2.3 T 38.1	7 6 3 7
French Conventional		3	25							FEB. 1947	180	FORM Q R 62.1 V 69.2 G 62.4 T 66.2	2 YRS. 31 37 54 40	UPPER B I 6.4 II 9.3 III 3.7 T 19.4	11 6 4 5 5
German Conventional		3	45							FEB. 1947	180	FORM P R 59.8 V 61.4 G 55.3 T 59.7	2 YRS. 25 25 23 23	UPPER A I 5.3 II 12.9 III 6.2 T 24.4	11-12 4 5 5 6
Spanish Conventional		3	43							FEB. 1947	180	FORM P R 70.2 V 70.6 G 70.7 T 72.5	2 YRS. 67 66 81 72	UPPER B I 10.2 II 7.5 III 5.3 T 23.0	11 5 6 4 6

COLLEGE J French	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	AURAL TENTH
X-1A	10	17						MAR. 1945	90	FORM O R 42.1 V 57.8 G 47.9 T 47.2	1 YR. 11 42 25 17	JUNE 1945	180	FORM O R 44.5 V 63.2 G 52.5 T 53.5	1 YR. 16 65 42 37	LOWER A I 6.1 II 12.2 III 14.4 T 32.7	1B 7 5 8 8
										R 36.8 V 58.3 G 49.3 T 48.1	4 44 29 19			R 41.6 V 63.8 G 55.0 T 53.8	10 67 53 38	I 6.9 II 12.9 III 15.3 T 35.1	8 7 8 8
X-1A.1*		9	DEC. 1945	250	FORM R R 52.4 V 60.4 G 57.3 T 57.2	UPPER B I 2.0 II 11.1 III 7.3 T 20.4	11 2 5 8 5	MAR. 1946	320	FORM O R 55.6 V 69.0 G 56.7 T 61.4	- - - -	JUNE 1946	390	FORM R R 65.8 V 66.1 G 61.6 T 66.0	2 YRS. 45 25 50 39	UPPER A I 11.3 II 18.2 III 3.9 T 33.4	11 6 5 7 7
1A	4	49	JAN. 1945	35	FORM O R 38.9 V 58.0 G 50.9 T 49.3			MAR. 1945	70	FORM O R 48.1 V 62.9 G 58.5 T 56.8	- - - -	JUNE 1945	105	FORM O R 56.6 V 69.5 G 62.8 T 63.6	1 YR. 58 84 81 75		
					R 37.6 V 56.6 G 50.6 T 48.1					R 48.0 V 62.1 G 59.4 T 56.6	- - - -			R 59.4 V 68.1 G 65.6 T 65.1	67 81 88 80		
1A.1*		17	DEC. 1945	140	FORM R R 71.2 V 65.1 G 64.1 T 66.9	UPPER B I 4.6 II 8.6 III 4.2 T 17.4	11 5 4 5 4	MAR. 1946	175	FORM O R 66.5 V 74.0 G 68.9 T 72.0	- - - -	JUNE 1946	210	FORM R R 79.1 V 73.2 G 67.7 T 75.3	2 YRS. 87 54 75 74	UPPER A I 10.4 II 18.5 III 3.4 T 32.3	11 6 5 6 6
X-1B	7											JUNE 1946	180	FORM R R 29.0 V 53.4 G 46.1 T 42.2	1 YR. 0 26 19 7	LOWER A I 4.1 II 11.5 III 4.8 T 20.4	1B 5 5 2 4
		20	DEC. 1946	240	FORM Q R 48.3 V 65.5 G 54.3 T 56.6	UPPER B♦ I 9.5 II 16.8 III 6.4 T 32.7	11 9 8 7 9	MAR. 1947	300	FORM O R 53.7 V 69.3 G 55.9 T 60.5	- - - -	JUNE 1947	360	FORM Q R 70.5 V 75.6 G 67.0 T 73.2	2 YRS. 63 63 73 67	UPPER A I 11.9 II 21.6 III 5.1 T 38.6	11 7 8 8 8
1B	4	45	DEC. 1945	35	FORM R R 37.0 V 51.6 G 54.8 T 47.7	LOWER B♦ I 6.0 II 11.4 III 7.7 T 25.1	- - - -	MAR. 1946	70	FORM O R 45.0 V 60.2 G 58.2 T 54.9	- - - -	JUNE 1946	105	FORM R R 57.7 V 61.9 G 58.8 T 60.4	1 YR. 61 60 64 65	LOWER A I 5.8 II 13.6 III 8.9 T 28.3	1B 7 7 4 7
X-1C.1	7	14				LOWER B♦ I 6.7 II 16.4 III 6.5 T 29.6	1B 6 7 4 7	MAR. 1947	120	FORM O R 36.1 V 63.6 G 47.4 T 48.9	1 YR. 4 66 23 25	JUNE 1947	180	FORM Q R 50.2 V 68.6 G 57.4 T 59.6	1 YR. 33 83 63 61	LOWER A I 6.7 II 16.4 III 14.0 T 37.1	1B 8 8 8 8
X-1C.2	7	14				I 9.9 II 16.1 III 8.1 T 34.1	8 7 5 8			R 47.5 V 62.8 G 58.4 T 56.7	24 63 67 50			R 56.9 V 68.2 G 66.8 T 65.4	59 82 91 81	I 7.8 II 17.7 III 13.9 T 39.4	9 8 8 8
X-2A#	8							MAR. 1945	H.S. +70	FORM O R 66.2 V 76.1 G 71.2 T 73.4	2 YRS. 46 65 86 67	JUNE 1945	H.S. +140	FORM O R 73.3 V 78.1 G 71.5 T 76.3	2 YRS. 73 73 87 78		
		11	DEC. 1945	H.S. +210	FORM R R 86.4 V 78.8 G 67.5 T 82.5	UPPER B I 10.0 II 17.0 III 9.0 T 36.0	12 6 5 7 7	MAR. 1946	H.S. +255	FORM O R 78.0 V 83.0 G 71.4 T 80.1	- - - -	JUNE 1946	H.S. +300	FORM R R 86.8 V 81.4 G 74.0 T 83.9	3 YRS. 88 50 83 78	UPPER A I 17.5 II 29.3 III 6.7 T 53.5	12 7 7 6 8

COLLEGE J French	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
2A	4	39	JAN. 1945	H.S. +35	FORM O			MAR. 1945	H.S. +70	FORM O		JUNE 1945	H.S. +105	FORM O	2 YRS.		
					R 60.5					R 66.5	-			R 74.1	75		
					V 72.1					V 74.3	-			V 79.4	77		
					G 62.7					G 67.3	-			G 68.8	79		
					T 66.7					T 71.4	-			T 76.5	79		
3⊕	3	19	DEC. 1945	H.S. +28	FORM R	UPPER B	12	MAR. 1946	H.S. +55	FORM O		JUNE 1946	H.S. +80	FORM R	3 YRS.	UPPER A	12
					R 88.0	I 7.7	4			R 74.7	-			R 91.0	94	I 13.2	4
					V 80.6	II 15.9	5			V 80.5	-			V 82.1	52	II 21.1	4
					G 71.8	III 7.1	5			G 69.5	-			G 72.1	77	III 4.9	5
					T 83.4	T 30.7	5			T 77.5	-			T 84.9	81	T 39.2	5
X-2B	7	8	DEC. 1945	H.S. +60	FORM R	UPPER B	11	MAR. 1946	H.S. +120	FORM O		JUNE 1946	H.S. +180	FORM R	2 YRS.	UPPER A	11
					R 77.4	I 6.1	6			R 71.4	-			R 88.2	98	I 13.1	7
					V 74.8	II 16.5	8			V 79.3	-			V 77.1	69	II 20.8	7
					G 72.5	III 6.1	7			G 71.4	-			G 69.8	81	III 5.2	8
					T 77.6	T 28.7	8			T 76.3	-			T 81.1	90	T 39.1	8
2B	4	58	DEC. 1945	H.S. +35	FORM R	UPPER B	11	MAR. 1946	H.S. +70	FORM O		JUNE 1946	H.S. +105	FORM R	2 YRS.	UPPER A	11
					R 71.5	I 5.0	5			R 65.7	-			R 78.6	86	I 10.4	6
					V 69.0	II 8.4	4			V 74.8	-			V 75.3	63	II 17.6	5
					G 64.8	III 4.3	5			G 66.7	-			G 68.7	78	III 3.6	7
					T 70.3	T 17.7	5			T 71.0	-			T 76.7	79	T 31.6	6
X-2C	7	6				UPPER B♦	11	MAR. 1947	H.S. +120	FORM O	2 YRS.	JUNE 1947	H.S. +160	FORM Q	2 YRS.	UPPER A	11
						I 10.7	9			R 63.7	36			R 75.8	80	I 14.2	8
						II 20.0	9			V 73.2	54			V 78.8	76	II 16.7	5
						III 9.0	9			G 65.7	67			G 68.7	80	III 6.7	9
						T 39.7	9			T 68.7	50			T 77.0	80	T 37.5	8

*This subgroup continued through two years.

♦This aural testing took place in March.

#In the second year, hours per week were 8 in first quarter, 5 in second and third quarters.

⊕Some of these students were taking another advanced French course concurrently.

COLLEGE J Spanish	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
X-1A.1	10	8						MAR. 1945	90	FORM O R 45.9 V 45.1 G 35.1 T 41.1	1 YR. 20 14 3 8	JUNE 1945	180	FORM O R 56.4 V 46.7 G 38.4 T 47.8	1 YR. 61 18 7 24	LOWER A I 8.5 II 12.1 III 12.0 T 32.6	1B 8 6 5 7
										R 54.4 V 52.4 G 36.2 T 47.4	52 37 4 23			R 60.6 V 54.8 G 40.7 T 52.3	76 47 11 41		
X-1A.2*		9	DEC. 1945	250	FORM O R 64.0 V 56.6 G 49.9 T 57.8	UPPER B I 13.1 II 13.0 III 4.4 T 30.5	11 8 8 3 7	MAR. 1946	320	FORM Q R 65.6 V 62.3 G 60.1 T 68.6	 - - - -	JUNE 1946	390	FORM O R 70.6 V 65.9 G 61.6 T 67.6	2 YRS. 69 46 46 55	UPPER A I 15.2 II 26.1 III 5.4 T 46.7	11 9 7 6 8
1A	4	84	JAN. 1945	35	FORM O R 45.5 V 48.3 G 43.3 T 45.2			MAR. 1945	70	FORM O R 56.2 V 55.5 G 51.5 T 54.6	 - - - -	JUNE 1945	105	FORM O R 62.7 V 62.0 G 56.3 T 61.4	1 YR. 82 75 66 77		
1A.1*					R 46.5 V 46.3 G 43.0 T 44.4					R 55.7 V 54.8 G 50.7 T 53.8	- - - -			R 62.8 V 61.8 G 56.2 T 61.2	82 74 66 76		
		30	DEC. 1945	140	FORM O R 66.6 V 67.8 G 56.9 T 65.3	UPPER B I 7.9 II 7.4 III 5.8 T 21.1	11 4 5 5 5	MAR. 1946	175	FORM Q R 68.9 V 70.9 G 66.5 T 70.7	 - - - -	JUNE 1946	210	FORM O R 72.2 V 73.9 G 68.9 T 73.8	2 YRS. 75 77 75 78	UPPER A I 9.2 II 24.3 III 4.7 T 38.2	11 5 6 5 6
X-1B	7		DEC. 1945	60	FORM O R 50.6 V 47.8 G 38.0 T 45.0	LOWER B♦ I 10.9 II 11.0 III 13.8 T 35.7	1B 8 6 5 7	MAR. 1946	120	FORM Q R 55.8 V 52.5 G 51.2 T 53.7	1 YR. 58 37 45 46	JUNE 1946	180	FORM O R 59.8 V 53.7 G 56.3 T 57.4	1 YR. 73 41 66 62	LOWER A I 8.4 II 12.4 III 14.5 T 35.3	1B 8 6 6 7
		25	DEC. 1946	240	FORM O R 64.3 V 56.7 G 63.7 T 62.8	UPPER B I 11.7 II 8.8 III 7.7 T 28.2	11 8 6 7 7	MAR. 1947	300	FORM Q R 66.4 V 61.2 G 64.9 T 65.6	 - - - -	JUNE 1947	360	FORM O R 68.9 V 66.9 G 67.3 T 69.6	2 YRS. 63 50 64 63	UPPER A I 11.2 II 23.5 III 6.2 T 40.9	11 6 6 7 7
1B	4	88	DEC. 1945	35	FORM O R 47.4 V 48.0 G 46.0 T 47.0			MAR. 1946	70	FORM Q R 59.7 V 57.1 G 58.7 T 59.5	 - - - -	JUNE 1946	105	FORM O R 64.8 V 61.1 G 62.6 T 64.2	1 YR. 88 72 86 85	LOWER A I 6.0 II 12.0 III 13.8 T 31.8	1B 6 6 6 7
X-1C	7	27				LOWER B♦ I 11.9 II 14.5 III 14.7 T 41.1	1B 9 8 5 8	MAR. 1947	120	FORM Q R 56.7 V 53.2 G 49.8 T 53.7	1 YR. 62 40 39 46	JUNE 1947	180	FORM O R 63.3 V 54.1 G 56.1 T 58.9	1 YR. 84 43 65 68	LOWER A I 8.2 II 12.0 III 15.0 T 35.2	1B 8 6 6 7
X-2A	8	23						MAR. 1945	H.S. +70	FORM O R 70.2 V 65.8 G 57.3 T 65.7	2 YRS. 68 45 29 46	JUNE 1945	H.S. +140	FORM O R 74.5 V 72.0 G 60.4 T 70.8	2 YRS. 82 70 42 68	UPPER A I 15.8 II 31.3 III 6.4 T 53.5	11 9 9 7 9
X-2A.1*#										R 71.2 V 66.2 G 59.7 T 67.1	71 47 38 53			R 76.2 V 76.1 G 64.1 T 74.0	87 82 56 79	I 17.2 II 31.1 III 7.2 T 55.5	9 9 8 9
		13	DEC. 1945	H.S. +210	FORM O R 84.0 V 78.3 G 64.6 T 76.7	UPPER B I 14.3 II 13.7 III 8.2 T 36.2	12 5 5 4 5	MAR. 1946	H.S. +255	FORM Q R 73.6 V 78.0 G 76.4 T 78.0	 - - - -	JUNE 1946	H.S. +300	FORM O R 80.3 V 82.4 G 73.8 T 81.5	3 YRS. 76 75 76 79	UPPER A I 19.2 II 32.3 III 7.7 T 59.2	12 9 9 9 9

COLLEGE J Spanish	HRS. PER WK.#	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
2A	4	107	JAN. 1945	H.S. +35	FORM O R 65.6 V 66.9 G 58.0 T 63.7			MAR. 1945	H.S. +70	FORM O R 70.3 V 71.0 G 61.2 T 69.0	 - - - -	JUNE 1945	H.S. +105	FORM O R 78.4 V 80.0 G 65.4 T 76.8	2 YRS. 91 91 62 86		
2A.1*		13			R 66.1 V 68.9 G 60.6 T 66.2					R 73.8 V 72.5 G 64.0 T 71.7	- - - -			R 78.8 V 81.7 G 65.0 T 77.6	91 93 60 87		
			DEC. 1945♦	H.S. +133	FORM O R 74.6 V 81.0 G 64.2 T 76.3	UPPER B I 12.9 II 13.3 III 6.8 T 33.0	12 4 4 4 4	MAR. 1946	H.S. +160	FORM Q R 77.2 V 85.5 G 72.8 T 81.2	 - - - -	JUNE 1946	H.S. +185	FORM O R 78.6 V 85.3 G 74.2 T 82.2	3 YRS. 70 83 78 80	UPPER A I 15.2 II 29.2 III 5.5 T 49.9	12 7 7 7 8
3	3	29	DEC. 1945	H.S. +28	FORM O R 71.8 V 73.8 G 56.6 T 69.1	UPPER B I 9.4 II 11.8 III 7.2 T 28.4	12 2 4 4 3	MAR. 1946	H.S. +55	FORM Q R 74.7 V 77.5 G 64.4 T 74.4	 - - - -	JUNE 1946	H.S. +80	FORM O R 73.6 V 79.7 G 66.8 T 75.5	3 YRS. 50 65 50 57	UPPER A I 12.1 II 26.5 III 4.5 T 43.1	12 5 6 5 6
X-2B⊕	7	27	DEC. 1945	H.S. +60	FORM O R 68.2 V 67.6 G 61.0 T 67.2	UPPER B I 10.8 II 9.8 III 8.2 T 28.8	11 5 7 7 7	MAR. 1946	H.S. +120	FORM Q R 71.2 V 70.4 G 71.6 T 73.1	 - - - -	JUNE 1946	H.S. +180	FORM O R 72.2 V 73.4 G 70.9 T 74.2	2 YRS. 75 74 81 80	UPPER A I 15.2 II 27.7 III 4.1 T 47.0	11 9 8 5 8
2B	4	68	DEC. 1945	H.S. +35	FORM O R 64.2 V 66.5 G 58.0 T 63.7	UPPER B I 6.5 II 6.0 III 5.9 T 18.4	11 3 5 5 3	MAR. 1946	H.S. +70	FORM Q R 68.8 V 70.3 G 65.9 T 70.2	 - - - -	JUNE 1946	H.S. +105	FORM O R 71.3 V 74.0 G 68.1 T 73.2	2 YRS. 71 77 72 76	UPPER A I 10.0 II 23.3 III 4.2 T 37.5	11 6 6 5 6
X-2C	7	38	DEC. 1946	H.S. +60	FORM O R 70.8 V 67.0 G 67.6 T 70.4	UPPER B I 12.0 II 11.3 III 7.7 T 31.0	11 8 7 7 7	MAR. 1947	H.S. +120	FORM Q R 72.4 V 73.0 G 72.5 T 74.8	 - - - -	JUNE 1947	H.S. +180	FORM O R 74.1 V 73.5 G 69.8 T 74.6	2 YRS. 80 75 78 80	UPPER A I 15.3 II 26.1 III 6.9 T 48.3	11 9 7 8 9

*This subgroup continued through two years.

♦This aural testing took place in March.

#In the second year, hours per week were 8 in first quarter, 5 in second and third quarters.

⊕Only testing in second year was Aural Upper B at H.S.+ 208 hrs. in Dec. 1946. Total mean: 40.3 (6th tenth).

COLLEGE J German	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
X-1	7	16				LOWER A* I 6.8 II 5.6 III 8.8 T 21.2	1B 6 4 5 6	MAR. 1946	120	FORM O R 49.5 V 51.1 G 49.6 T 50.3	JUNE 1946	160	FORM Q R 54.2 V 55.5 G 54.1 T 55.0	1 YR. 42 46 55 48	LOWER A I 9.9 II 7.5 III 11.4 T 28.8	1B 8 6 6 7
1	4	96				LOWER A* I 4.8 II 4.9 III 7.3 T 17.0	1A 6 6 6 6	MAR. 1946	70	FORM O R 53.4 V 57.4 G 47.3 T 52.9	JUNE 1946	105	FORM Q R 57.3 V 60.2 G 49.7 T 56.1	1 YR. 54 66 38 52	LOWER A I 6.8 II 5.6 III 8.8 T 21.2	1B 6 4 5 6
X-2	7	9	DEC. 1945	H.S. +60	FORM Q R 59.6 V 58.7 G 59.1 T 59.9	UPPER A I 5.6 II 20.6 III 7.0 T 33.2	11-12 5 8 6 8	MAR. 1946	H.S. +120	FORM O R 62.9 V 65.8 G 54.8 T 62.2	JUNE 1946	H.S. +180	FORM Q R 65.7 V 65.8 G 62.4 T 66.0	2 YRS. 44 43 50 45	UPPER A I 13.7 II 29.6 III 11.4 T 54.7♦	11-12 9 9 9 9
2	4	54	DEC. 1945	H.S. +35	FORM Q R 58.1 V 58.0 G 55.9 T 58.0	UPPER A I 4.5 II 13.4 III 5.3 T 23.2	11-12 4 5 5 5	MAR. 1946	H.S. +70	FORM O R 60.8 V 62.1 G 53.1 T 59.4	JUNE 1946	H.S. +105	FORM Q R 67.0 V 64.6 G 60.6 T 65.5	2 YRS. 50 39 43 43	UPPER A I 5.3 II 15.8 III 6.3 T 27.4	11-12 5 6 5 7
3	3	8	DEC. 1945	H.S. +28	FORM Q R 73.3 V 71.0 G 66.8 T 72.3	UPPER A I 7.2 II 15.8 III 5.2 T .28.2	11-12 7 6 5 7	MAR. 1946	H.S. +55	FORM O R 74.7 V 73.7 G 62.3 T 72.2	JUNE 1946	H.S. +80	FORM Q R 74.0 V 71.8 G 69.4 T 73.8	3 YRS. 51 32 61 48	UPPER A I 9.9 II 22.1 III 9.3 T 41.3	11-12 8 8 8 8

*This aural testing took place in March.

♦This group had already taken Aural Upper A in March as well as in December. Total mean in March: 48.4 (9th tenth).

COLLEGE B Experimental	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	%ILE	AURAL MEAN	TENTH
French	1A	9	5	FEB. 1946	150	FORM R R 52.2 V 58.8 G 57.4 T 56.8	1 YR. 40 47 63 50	LOWER B I 10.6 II 17.8 III 13.6 T 42.0	2 9 8 7 9						
Spanish	1A	9	8	FEB. 1946	150	FORM Q R 65.1 V 65.3 G 65.0 T 66.6	1 YR. 88 84 91 90	LOWER B I 13.3 II 15.0 III 20.5 T 48.8	1B 9 8 8 9						
French	1B	9	7	MAR. 1947	180	FORM R R 50.1 V 62.1 G 59.4 T 57.6	1 YR. 32 60 70 53	LOWER A I 7.9 II 11.1 III 13.9 T 32.9	2 7 3 6 6						
Spanish	1B	9	12	MAR. 1947	185	FORM Q R 64.4 V 63.5 G 69.1 T 67.3	1 YR. 87 79 96 91	LOWER A I 13.1 II 16.3 III 15.0 T 44.4	1B 9 9 6 9						
French	2A	6	3	OCT. 1945	H.S. +30	FORM R R 63.0 V 69.7 G 68.3 T 68.7	1 YR. 79 85 93 89	UPPER A I 8.0 II 15.0 III 5.0 T 28.0	11 5 4 8 5	APR. 1946	H.S. +140	FORM R R 71.0 V 73.0 G 74.7 T 75.0	3 YRS. 35 18 85 44	UPPER B I 5.7 II 12.0 III 5.7 T 23.3	12 2 4 4 3
Spanish	2A	6	5	NOV. 1945	H.S. +30	FORM Q R 73.6 V 73.4 G 76.8 T 77.0	1 YR. 98 97 99 99	UPPER A I 15.8 II 35.6 III 10.2 T 61.6	11 9 10 10 10	APR. 1946	H.S. +140	FORM Q R 79.8 V 89.0 G 79.4 T 85.6	3 YRS. 75 91 90 89	UPPER A I 21.4 II 38.6 III 9.2 T 69.2	12 10 10 9 10
French	2B	6	11							MAR. 1947	H.S. +120	FORM R R 80.7 V 79.2 G 71.1 T 77.2	2 YRS. 90 76 86 81	UPPER A I 16.3 II 28.8 III 4.4 T 49.5	11 9 9 7 9
Spanish	2B	6	12							MAR. 1947	H.S. +120	FORM Q R 68.4 V 72.4 G 75.4 T 74.4	2 YRS. 61 72 91 80	UPPER A I 11.9 II 26.2 III 5.3 T 43.4	11 8 7 6 7
Spanish	3A	4	9	OCT. 1945	280	FORM Q R 75.4 V 85.9 G 76.9 T 82.1	2 YRS. 84 97 93 95	UPPER A I 16.1 II 26.1 III 5.2 T 47.4	12 8 6 6 7	APR. 1946	350	FORM Q R 79.4 V 92.3 G 79.8 T 86.9	3 YRS. 73 95 91 91	UPPER A I 18.4 II 35.1 III 8.6 T 62.1	12 9 9 9 10
Spanish	3B	4	4*							MAR. 1947	350	FORM Q R 74.0 V 93.0 G 79.0 T 85.0	3 YRS. 52 96 90 87	UPPER A I 11.8 II 27.5 III 4.8 T 44.1	12 5 7 6 7
French	4	4	3	OCT. 1945	195	FORM R R 91.3 V 82.0 G 76.7 T 87.0	3 YRS. 94 52 89 86	UPPER A I 21.0 II 35.7 III 8.3 T 65.0♦	12 9 9 7 9	APR. 1946	280	FORM R R 98.3 V 86.7 G 70.0 T 88.7	4 YRS. 99 53 58 80	UPPER B I 10.3 II 20.3 III 8.7 T 39.3	12 6 7 7 7
Spanish	4	4	2	OCT. 1945	195	FORM Q R 71.0 V 90.0 G 81.5 T 83.5	3 YRS. 40 92 94 85	UPPER A I 14.5 II 26.0 III 6.0 T 46.5	12 7 6 7 7	APR. 1946	280	FORM Q R 82.5 V 96.0 G 78.5 T 89.0	4 YRS. 63 93 79 85	UPPER A I 21.0 II 35.0 III 8.5 T 64.5	12 10 9 9 10
RUSSIAN	1	9	20	MAR. 1947	180	INV. R 18.0 V 12.7 G 16.0 T 46.7	TENTH 6 5 5 5	LOWER A I 9.3 II 13.1 III 14.0 T 36.4	 6 5 4 5						
	3A	6	6	OCT. 1945	290			LOWER A I 13.2 II 16.7 III 24.5 T 54.4	 7 7 8 8	APR. 1946	400			UPPER A I 10.7 II 22.0 III 10.5 T 43.2	 5 5 7 6
	3B	6	4							MAR. 1947	400	INV. R 19.5 V 17.5 G 15.5 T 52.5	TENTH 6 8 5 6	UPPER A I 8.8 II 18.5 III 1.5 T 28.8	 4 4 1 2

*Two of the four failed to take the Co-operative test.

♦Six months earlier this group had taken a tryout form of Aural Upper A.

6

COLLEGE C June 1947	GROUP	HRS. PER WK.	N	TOTAL HRS.	CO-OPERATIVE		AURAL		Further Aural Notes
					MEAN	%ILE	MEAN	TENTH	
Spanish	2	3	41	135	FORM O	2 YRS.	LOWER B	2	Twenty-one of these (*30.2) had taken Aural Lower A at 90 hrs. in Jan. 1947. Total mean: 27.8 (5th tenth).
					R 60.7	31	I 8.7	4	
					V 61.8	30	II 10.9	4	
					G 61.2	44	III 13.1	4	
					T 62.6	35	T 32.7*	4	
French	3	3	110	180	FORM P	2 YRS.	UPPER B	11	
					R 69.6	61	I 6.0	6	
					V 74.0	57	II 12.8	6	
					G 65.7	69	III 4.8	6	
					T 72.1	63	T 23.6	7	
Spanish	3	3	118	180	FORM O	2 YRS.	UPPER B	11	Eighty-four of these (*25.9) had taken Aural Lower A at 135 hrs. in Jan. 1947. Total mean: 37.5 (6th tenth).
					R 67.9	59	I 10.7	5	
					V 66.6	50	II 9.0	6	
					G 65.9	64	III 6.7	5	
					T 68.5	57	T 26.4*	6	
French	x♦	9	13	270	FORM P	2 YRS.	UPPER B	11	
					R 75.7	80	I 12.2	10	
					V 73.5	57	II 16.5	8	
					G 66.5	69	III 6.2	7	
					T 74.2	70	T 34.9	10	
Spanish	x♦	9	15	270	FORM O	2 YRS.	UPPER B	11	Fourteen of these (*40.6) had taken Aural Lower A at 135 hours in Jan. 1947. Total mean: 41.8 (9th tenth).
					R 70.6	70	I 14.1	9	
					V 67.7	54	II 17.1	10	
					G 72.1	84	III 8.2	7	
					T 72.2	72	T 39.4*	10	
Spanish	4	3	30	225	FORM O	3 YRS.			This group took no aural test in June 1947, but twenty of them had taken Aural Upper A at 180 hrs. in Jan. 1947. Total mean: 31.2 (4th tenth).
					R 67.1	26			
					V 69.9	28			
					G 60.4	26			
					T 67.5	26			
French	5	3	47	OVER 225	FORM P	3 YRS.	UPPER B	12	
					R 79.3	66	I 10.9	6	
					V 81.5	50	II 16.3	5	
					G 74.0	83	III 6.8	5	
					T 81.1	68	T 34.0	6	
Spanish	5	3	23	OVER 225	FORM O	3 YRS.	UPPER B	12	Twelve of these (*37.2) had taken Aural Upper A at 225 hours in Jan. 1947. Total mean: 37.7 (5th tenth).
					R 73.4	50	I 13.0	4	
					V 74.9	47	II 17.8	7	
					G 70.9	67	III 8.8	5	
					T 75.2	56	T 39.6*	6	
Spanish	6	3	4	OVER 225	FORM O	3 YRS.	UPPER B	12	
					R 79.5	73	I 16.3	7	
					V 85.8	85	II 19.3	7	
					G 78.0	88	III 11.3	8	
					T 84.0	85	T 46.8	8	

♦X designates Experimental Group.

COLLEGE K	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH	DATE	TOTAL HRS.	MEAN	%ILE	MEAN	TENTH
FRENCH	X-1	15	16	DEC. 1945	180	FORM R R 41.1 V 55.2 G 53.4 T 49.9	1 YR. 10 32 45 25	LOWER A I 2.4 II 8.4 III 8.9 T 19.8	1B 3 3 4 4	MAR. 1946	360	FORM R R 65.1 V 67.6 G 60.2 T 65.6	1 YR. 84 80 73 82	UPPER A I 4.8 II 15.9 III 2.6 T 23.3	11 3 5 6 3
	1	5	50	JUNE 1946	180	FORM R R 68.0 V 70.4 G 67.1 T 71.1	1 YR. 90 86 91 93	LOWER A I 2.8 II 11.9 III 9.7 T 24.4	1B 4 5 5 [illegible]						
GERMAN	X-1	10	14	MAR. 1946	240	FORM Q R 51.1 V 47.5 G 52.6 T 50.3	1 YR. 31 18 51 29	LOWER A I 7.4 II 6.6 III 8.6 T 22.6	1B 6 5 5 6	JUNE 1946	360	FORM Q R 56.1 V 57.0 G 56.9 T 57.1	1 YR. 49 53 67 57	UPPER A I 6.6 II 18.1 III 7.9 T 32.6	11-12 7 7 7 8
	1	5	49	MAR. 1946	120	FORM Q R 52.9 V 50.5 G 55.0 T 53.0	1 YR. 38 28 59 40	LOWER A I 4.8 III 4.4 III 7.8 T 17.0	1B 5 3 5 4						
SPANISH	X-1	10	26	DEC. 1945	120	FORM Q R 45.5 V 40.5 G 42.4 T 42.2	1 YR. 19 6 14 9	LOWER A I 3.7 II 8.0 III 14.5 T 26.2	1B 4 3 6 4	MAR. 1946	240	FORM Q R 52.3 V 46.7 G 55.3 T 51.7	1 YR. 43 19 61 40	UPPER A I 6.6 II 18.4 III 4.1 T 29.2	11 4 4 5 4
	1	5	78	MAY 1946	180	FORM Q R 64.7 V 65.0 G 64.4 T 66.5	1 YR. 88 84 90 90	LOWER A I 5.2 II 11.2 III 13.9 T 30.3	1B 5 6 6 5						
	X-2	10	11							DEC. 1945	480	FORM Q R 67.6 V 66.4 G 68.7 T 67.5	2 YRS. 58 48 76 56	UPPER A I 13.5 II 26.8 III 3.5 T 43.8	11 9 7 5 8

COLLEGE L	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
French	X-1A	5	184	APR. 1946	110			LOWER A I 4.5 II 12.2 III 10.3 T 27.0	1B 5 6 5 6
French	1A	3	53	JULY 1946	110			LOWER A I 4.2 II 13.0 III 10.6 T 27.8	1B 5 6 6 6
Spanish	X-1A	5	15	JAN. 1946	110			LOWER A I 8.0 II 12.4 III 16.1 T 36.5	1B 8 6 7 8
Spanish	1A	3	132	JAN. 1946	130			LOWER A I 5.5 II 10.1 III 11.7 T 27.3	1B 6 5 5 5
French	X-1B	5	140	MAR. 1947	110	FORM R* R 52.7 V 56.7 G 52.4 T 54.3	1 YR. 43 39 40 39	LOWER A I 4.9 II 10.0 III 9.6 T 24.5	1B 6 4 5 5
French	1B	5	149	MAR. 1947	110	FORM R* R 56.7 V 59.7 G 57.6 T 58.7	1 YR. 57 52 65 59	LOWER A I 4.3 II 10.5 III 10.5 T 25.3	1B 5 4 5 5
Spanish	X-1B	5	263	MAR. 1947	110	FORM O OR P* R 60.2 V 53.7 G 53.7 T 56.6	1 YR. 74 43 57 60	LOWER A I 6.7 II 11.2 III 15.3 T 33.2	1B 7 6 6 7
Spanish	1B	5	287	MAR. 1947	110	FORM O OR P* R 59.0 V 52.9 G 55.3 T 56.4	1 YR. 71 39 61 56	LOWER A I 5.9 II 10.7 III 14.9 T 31.5	1B 6 6 6 7

*Elementary Form

COLLEGE M	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	MEAN	TENTH	DATE	TOTAL HRS.	CO-OPERATIVE MEAN	CO-OPERATIVE %ILE	AURAL MEAN	AURAL TENTH
German	X-1	9	39	MAR. 1946	200	LOWER A I 14.1 II 13.4 III 14.6 T 42.2	2 9 8 7 9	MAY 1946	270	FORM Q R 71.4 V 69.2 G 78.3 T 75.3	2 YRS. 66 57 94 79	UPPER A I 9.5 II 24.2 III 7.6 T 41.2	11-12 8 9 7 9

COLLEGE N	GROUP	HRS. PER WK.	N	DATE	TOTAL HRS.	CO-OPERATIVE		AURAL	
						MEAN	%ILE	MEAN	TENTH
FRENCH	X-1A	6	25	MAY 1946	240	FORM R R 57.7 V 63.6 G 58.2 T 60.8	1 YR. 62 68 65 67	LOWER A I 5.7 II 11.1 III 8.0 T 24.8*	1B 7 5 4 5
	1A	5	93	MAY 1946	150	FORM R R 63.7 V 63.5 G 63.3 T 64.9	1 YR. 81 68 82 80	LOWER A I 3.2 II 10.6 III 5.4 T 19.1	1B 4 5 3 3
	X-1B	6	15	MAY 1947	240			LOWER A I 4.5 II 13.2 III 8.7 T 26.4	1B 5 6 5 6
	1B.1	5	219	MAY 1947	150			LOWER A I 2.8 II 9.2 III 8.2 T 20.2	1B 4 4 4 4
	1B.2	5	112	MAY 1947	150			LOWER A I 2.9 II 11.2 III 8.6 T 22.7	1B 4 5 5 5

*This testing took place in April 1946 at 200 hrs. The group was retested on same form in May (240 hrs.) and made a total mean of 31.6 (8th tenth).

HIGH SCHOOLS	LANGUAGE	HRS. PER WK.	N	DATE	AURAL	
					MEAN	TENTH
HIGH SCHOOL B	French	5	10	JAN. 1947	LOWER A	OB
					I 4.1	6
					II 9.7	6
					III 6.2	5
					T 20.0	6
	German	5	10	JAN. 1947	LOWER A	OB
					I 10.6	9
					II 7.1	8
					III 14.8	9
					T 32.5	9
	Spanish	5	10	JAN. 1947	LOWER A	OB
					I 8.7	10
					II 12.7	10
					III 13.8	8
					T 35.2	10
HIGH SCHOOL C	French	20	20	AUG. 1945	LOWER A	OB
					I 6.6	9
					II 8.6	5
					III 6.6	6
					T 21.8	7
	Spanish	20	25	AUG. 1945	LOWER A	OB
					I 7.4	8
					II 10.0	8
					III 15.1	9
					T 32.6	10
HIGH SCHOOL D	French 1a	5	9	JAN. 1946	LOWER A	OA
					T 21.1	7
				JUNE 1946	LOWER B	
					T 20.2	-
	French 1b	5	8	JAN. 1946	LOWER A	OA
					T 14.5	4
				JUNE 1946	LOWER B	
					T 13.4	-

Appendix C

The Reliabilities and Internal Correlations of the Aural Tests

The following figures are usually based on a sample of 200 papers. This sample contains a proportional representation of the institutions furnishing students for any particular group, but a random selection from the papers of each school. When fewer than 200 papers were available by the time these computations were made, all the papers were used and the size of the group has been indicated. The students on whose scores these figures are based were all conventionally taught (non-experimental) except in the case of the Russian tests, where the very small number of cases made it desirable to include experimental students in order to secure the largest possible body of data.

The reliabilities were computed by the use of odd-even split-halves, with the coefficient corrected by the Spearman-Brown formula. It is worth noting that, other things (such as the size of the samples) being equal, the reliability coefficients for any one test tend to increase along with greater training on the part of the students tested.

FRENCH AURAL TEST — LOWER LEVEL FORM A

College students with no high school French who had had 90-130 hours of college elementary instruction

	I	II	III	Tot.
I	.65			
II	.42	.77		
III	.43	.48	.73	
Tot.	.69	.79	.84	.86

College students taking intermediate or advanced college courses (N=196)

	I	II	III	Tot.
I	.65			
II	.48	.75		
III	.52	.41	.76	
Tot.	.78	.76	.83	.88

FRENCH AURAL TEST — UPPER LEVEL FORM A

College students with no high school French who were tested near the end of an intermediate course (N=89)

	I	II	III	Tot.
I	.59			
II	.59	.76		
III	.38	.54	.50	
Tot.	.81	.92	.66	.85

College students with various amounts of high school French who were tested in first part of intermediate course

	I	II	III	Tot.
I	.76			
II	.56	.81		
III	.34	.41	.46	
Tot.	.82	.90	.56	.88

College students with various amounts of high school French who were tested near the end of an intermediate course (N=181)

	I	II	III	Tot.
I	.75			
II	.56	.71		
III	.44	.36	.30	
Tot.	.84	.88	.60	.82

College students taking an advanced course (N=141)

	I	II	III	Tot.
I	.83			
II	.69	.86		
III	.51	.63	.71	
Tot.	.85	.94	.75	.92

GERMAN AURAL TEST — LOWER LEVEL FORM A

College students with no high school German who had had less than 90 hours of college elementary instruction

	I	II	III	Tot.
I	.70			
II	.60	.53		
III	.48	.50	.70	
Tot.	.80	.79	.84	.84

College students with no high school German who had had 90-130 hours of college elementary instruction

	I	II	III	Tot.
I	.76			
II	.68	.70		
III	.67	.64	.84	
Tot.	.86	.84	.90	.95

College students taking intermediate or advanced college courses (N=152)

	I	II	III	Tot.
I	.76			
II	.62	.74		
III	.59	.62	.82	
Tot.	.84	.84	.88	.89

GERMAN AURAL TEST — UPPER LEVEL FORM A

College students taking intermediate or advanced college courses (N=83)

	I	II	III	Tot.
I	.58			
III	.60	.86		
III	.46	.54	.54	
Tot.	.81	.92	.70	.79

RUSSIAN AURAL TEST — LOWER LEVEL FORM A

College students taking elementary courses (N=155)

	I	II	III	Tot.
I	.80			
II	.77	.79		
III	.74	.68	.87	
Tot.	.91	.88	.92	.94

RUSSIAN AURAL TEST — UPPER LEVEL FORM A

College students taking intermediate or advanced courses (N=60)

	I	II	III	Tot.
I	.74			
II	.54	.83		
III	.52	.72	.73	
Tot.	.78	.92	.84	.88

SPANISH AURAL TEST — LOWER LEVEL FORM A

College students with no high school Spanish who had had 90-130 hours of college elementary instruction

	I	II	III	Tot.
I	.45			
II	.52	.59		
III	.39	.36	.78	
Tot.	.74	.75	.82	.82

College students with various amounts of high school Spanish taking an intermediate college course

	I	II	III	Tot.
II	.41			
II	.54	.65		
III	.47	.45	.79	
Tot.	.77	.76	.85	.84

College students taking intermediate or advanced college courses

	I	II	III	Tot.
I	.67			
II	.66	.73		
III	.59	.59	.80	
Tot.	.85	.84	.87	.90

SPANISH AURAL TEST — UPPER LEVEL FORM A

College students with no high school Spanish who were tested near the end of an intermediate course

	I	II	III	Tot.
I	.50			
II	.53	.72		
III	.32	.34	.57	
Tot.	.78	.89	.56	.76

College students with various amounts of high school Spanish who were tested in first part of an intermediate course

	I	II	III	Tot.
I	.28			
II	.31	.68		
III	.07	.20	.51	
Tot.	.61	.89	.45	.73

College students with various amounts of high school Spanish who were tested near the end of an intermediate course

	I	II	III	Tot.
I	.61			
II	.55	.70		
III	.41	.42	.40	
Tot.	.81	.89	.63	.78

College students taking an advanced course

	I	II	III	Tot.
I	.80			
II	.61	.72		
III	.47	.32	.53	
Tot.	.85	.89	.65	.84

Printed in the United States of America